Undersampled Dynamic Magnetic Resonance Imaging using Kernel Principal Component Analysis

Yanhua Wang, and Leslie Ying, *IEEE Senior Member*

***Abstract*—Compressed sensing (CS) is a promising approach to accelerate dynamic magnetic resonance imaging (MRI). Most existing CS methods employ linear sparsifying transforms. The recent developments in non-linear or kernel-based sparse representations have been shown to outperform the linear transforms. In this paper, we present an iterative non-linear CS dynamic MRI reconstruction framework that uses the kernel principal component analysis (KPCA) to exploit the sparseness of the dynamic image sequence in the feature space. Specifically, we apply KPCA to represent the temporal profiles of each spatial location and reconstruct the images through a modified pre-image problem. The underlying optimization algorithm is based on variable splitting and fixed-point iteration method. Simulation results show that the proposed method outperforms conventional CS method in terms of aliasing artifact reduction and kinetic information preservation.**

I. Introduction

Dynamic magnetic resonance imaging (MRI) produces a series of images characterizing certain kinetic information of tissues. For Cartesian trajectories, the long data acquisition time limits the achievable spatiotemporal resolution. Reducing the amount of acquired data without degrading image quality is a possible approach to accelerate the imaging speed and thus improve the spatiotemporal resolution. A lot of efforts have been made to reconstruct the image sequence from undersampled measurements by exploiting either spatial or temporal correlations [1]-[7]. Among various approaches, compressed sensing (CS), which utilizes the sparseness of the image sequence, has demonstrated great potentials in image reconstruction from randomly undersampled k-space data [8]-[13]. One of the keys to the success of CS reconstruction lies in choosing a certain sparsifying transform that can make the image sequence as sparse as possible. Since the dynamic image sequence shows high correlations between adjacent frames, typical transforms employed by existing CS dynamic imaging methods are mostly applied along the temporal directions, such as Fourier transform [8]-[10][12], principal component analysis (PCA) [10], and dictionary learning [11][13]. These transforms are performed in the original linear space, which might not effectively represent the non-linear features of the temporal variations.

In recent years, kernel-based sparse representation has attracted a lot of attention. It has been applied in various applications, such as classification and recognition, and has demonstrated promising superiorities over linear sparse representations [14]-[17]. The basic concept of kernel sparse representation is constructing a non-linear mapping from the original signal space to a high dimensional feature space and looking for the sparse representation of the mapped signal. Kernel sparse representation has been incorporated into the CS reconstruction framework [18][19]. The kernel trick is used to extend both the measuring process and sparse representations to the high dimensional feature space. For CS dynamic imaging, a kernel-based method was recently proposed [20]. It explicitly maps the acquired data to a higher dimensional space and then performs CS reconstruction in the feature space. The desired image sequence is finally obtained by solving a pre-image problem. This method achieves better reconstruction qualities than conventional CS, but it requires the explicit knowledge of the mapping, which is not available for certain kernel functions [21].

In this paper, we propose a non-linear CS dynamic MRI reconstruction method that employs kernel sparse representation as regularizations. Specifically, we apply kernel PCA (KPCA) to represent the signal in the feature space. The resulting highly non-linear optimization problem is solved by variable splitting and fixed-point iteration method. Retrospective simulation results show that the proposed method is capable of reducing aliasing artifacts and preserving temporal features compared to the conventional method using linear PCA as the sparsifying transform.

This paper is structured as follows. In Section II, the basics of kernel method and KPCA are briefly described. Then, the proposed KPCA-based CS dynamic MRI model is introduced. Finally, the optimization algorithm is detailed. Section III shows the simulation results using an arterial spin labeled (ASL) perfusion data. The paper is concluded in Section IV.

II. Theory

A. Kernel Method and Kernel PCA

In a nutshell, kernel method aims to exploit the non-linear features of the data and develop the non-linear version of a given linear algorithm [21]. In kernel method, the original data is first transformed to a higher dimensional space using a

This work was supported in part by the National Science Foundation CBET-1265612.

Yanhua Wang is with the Department of Biomedical Engineering and the Department of Electrical Engineering, University at Buffalo, The State University of New York, Buffalo, NY 14260 USA (e-mail: wyhlucky@gmail.com).

Leslie Ying is with the Department of Biomedical Engineering and the Department of Electrical Engineering, University at Buffalo, The State University of New York, Buffalo, NY 14260 USA (phone: 716-645-1609; e-mail: leiying@buffalo.edu).

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

feature mapping function $\phi: \mathcal{X} \to \mathcal{H}$, where $\mathcal{X}$ denotes the original data space and $\mathcal{H}$ stands for the feature space. Then, the linear algorithm is performed in the feature space. For certain applications such as classification and recognition, the results from the feature space can be directly adopted. However, for reconstruction purposes, we have to transform the data back to the original space by solving a pre-imaging problem [21][23].

The dimension of feature space is always much higher than the original space. Performing linear algorithms on such high dimensional data may result in unaffordable computational complexities. Furthermore, the dimension of the feature space can be infinite in some cases, making it impossible to execute the algorithm directly. However, when the linear algorithms only need to compute inner products, the kernel trick can be applied to compute the inner product in the feature space using a kernel function without the explicit knowledge of the mapping. The kernel function is a symmetric function of two variables defined by

$$k(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle, \tag{1}$$

where $\mathbf{x}_i \in \mathcal{X}$, $\phi(\mathbf{x}_i) \in \mathcal{H}$. Typical kernel functions include the radial basis function (RBF) $k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-c\left\|\mathbf{x}_i - \mathbf{x}_j\right\|_2^2\right)$ with $c > 0$ controlling the width of the RBF, and polynomial function $k(\mathbf{x}_i, \mathbf{x}_j) = (\langle \mathbf{x}_i, \mathbf{x}_j \rangle + a)^b$, where a is a scalar and b is the degree of polynomial.

As a widely used dimensionality reduction method, PCA projects high dimensional data to a low dimensional sub-space. The projected data can be viewed as a sparse representation of the original data. KPCA is a non-linear version of PCA which performs the projection in the feature space [22]. Different from PCA, KPCA does not obtain the principal components (PCs) explicitly in the feature space. Since each PC can be represented by a linear combination of training signals, KPCA computes the representation coefficients of PCs instead. Given a set of L training signals $\mathbf{p}_l$ $(l = 1, 2, \cdots, L)$, a PC in feature space can be expressed by $\mathbf{V} = \sum_{l=1}^{L} \alpha_l \bar{\phi}(\mathbf{p}_l)$, where $\bar{\phi}(\mathbf{p}_l) = \phi(\mathbf{p}_l) - \sum_{l=1}^{L} \phi(\mathbf{p}_l) \big/ L$ and α_l $(l = 1, 2, \cdots, L)$ are the representation coefficients. Define an $L \times L$ kernel matrix

$$\mathbf{K}_\mathbf{p} = \begin{bmatrix} k(\mathbf{p}_1, \mathbf{p}_1) & k(\mathbf{p}_1, \mathbf{p}_2) & \cdots & k(\mathbf{p}_1, \mathbf{p}_L) \\ k(\mathbf{p}_2, \mathbf{p}_1) & k(\mathbf{p}_2, \mathbf{p}_2) & \cdots & k(\mathbf{p}_2, \mathbf{p}_L) \\ \vdots & \vdots & \ddots & \vdots \\ k(\mathbf{p}_L, \mathbf{p}_1) & k(\mathbf{p}_L, \mathbf{p}_2) & \cdots & k(\mathbf{p}_L, \mathbf{p}_L) \end{bmatrix}, \tag{2}$$

and the centered kernel matrix

$$\mathbf{K}_\mathbf{p}^C = \mathbf{K}_\mathbf{p} - \mathbf{1}_L \mathbf{K}_\mathbf{p} - \mathbf{K}_\mathbf{p} \mathbf{1}_L + \mathbf{1}_L \mathbf{K}_\mathbf{p} \mathbf{1}_L, \tag{3}$$

where $\mathbf{1}_L$ is an $L \times L$ matrix with all entries equal to $1/L$. The representation coefficients of PCs are obtained by solving the following eigenvalue problem

$$\mathbf{K}_\mathbf{p}^C \boldsymbol{\alpha} = \lambda \boldsymbol{\alpha}, \tag{4}$$

where $\sqrt{\lambda}$ is the length of the PC, $\boldsymbol{\alpha} = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_L]^T$.

Once the PCs are obtained, we can compute the projection of a test signal on a certain PC. For a test signal $\mathbf{x}$, define the kernel vector

$$\mathbf{k}_\mathbf{xp} = [k(\mathbf{p}_1, \mathbf{x}) \quad k(\mathbf{p}_2, \mathbf{x}) \quad \cdots \quad k(\mathbf{p}_L, \mathbf{x})]^T, \tag{5}$$

and the centered kernel vector with the entries calculated by

$$\begin{aligned} \mathbf{k}_\mathbf{xp}^C(l) = \mathbf{k}_\mathbf{xp}(l) - \frac{1}{L}\sum_{i=1}^{L} \mathbf{K}_\mathbf{p}(l, i) - \frac{1}{L}\sum_{i=1}^{L} \mathbf{k}_\mathbf{xp}(i) \\ + \frac{1}{L^2}\sum_{k=1}^{L}\sum_{i=1}^{L} \mathbf{K}_\mathbf{p}(k, i). \end{aligned} \tag{6}$$

The projection of centered $\phi(\mathbf{x})$, namely $\bar{\phi}(\mathbf{x})$, onto the k-th PC is computed by $\beta_k = (\mathbf{k}_\mathbf{xp}^C)^T \boldsymbol{\alpha}^k$. Similar to PCA, $\bar{\phi}(\mathbf{x})$ can be approximated by the first K largest PCs in the feature space

$$\bar{\phi}(\mathbf{x}) \approx \sum_{k=1}^{K} \beta_k \mathbf{V}^k = \sum_{k=1}^{K} \beta_k \sum_{l=1}^{L} \alpha_l^k \bar{\phi}(\mathbf{p}_l) = \sum_{l=1}^{L} \bar{\gamma}_l \bar{\phi}(\mathbf{p}_l), \tag{7}$$

where $\bar{\gamma}_l = \sum_{k=1}^{K} \beta_k \alpha_l^k$. Then, $\phi(\mathbf{x})$ is obtained by

$$\phi(\mathbf{x}) = \bar{\phi}(\mathbf{x}) + \frac{1}{L}\sum_{l=1}^{L} \phi(\mathbf{p}_l) \approx \sum_{l=1}^{L} \gamma_l \phi(\mathbf{p}_l), \tag{8}$$

where $\gamma_l = \bar{\gamma}_l + (1 - \sum_{l=1}^{L} \bar{\gamma}_l) \big/ L$. The above formulation (8) provides a sparse approximation of $\phi(\mathbf{x})$ and the sparsity is controlled by the number of PCs used for approximation.

B. Proposed Method

As shown above, KPCA offers a non-linear approach for sparse representation. In this paper, we aim to use KPCA to represent the temporal profiles of a dynamic image sequence. Let $\mathbf{x}$ be the vectorized image sequence. Define $\mathbf{R}_m$ as the operator that extracts the temporal curve at the m-th spatial location and $\mathbf{x}_m = \mathbf{R}_m \mathbf{x}$. We propose the following two-step iterative reconstruction scheme.

- Step 1: sparse approximation

  Using KPCA, $\phi(\mathbf{x}_m)$ can be approximated by

$$\phi(\mathbf{x}_m) \approx \sum_{l=1}^{L} \gamma_l^m \phi(\mathbf{p}_l). \tag{9}$$

  Here, the training signals $\mathbf{p}_l$ represent the temporal variation. They are chosen from the results of previous iteration. The number of PCs used to compute γ_l^m controls the sparsity level of $\phi(\mathbf{x}_m)$.

- Step 2: image reconstruction

  Once γ_l^m is obtained, $\mathbf{x}$ is reconstructed by

$$\min_{\mathbf{x}} \frac{1}{2}\left\|\mathbf{y}-\mathbf{F}_u\mathbf{x}\right\|_2^2+\eta\sum_m\left\|\phi(\mathbf{x}_m)-\sum_{l=1}^{L}\gamma_l^m\phi(\mathbf{p}_l)\right\|_2^2, \quad (10)$$

where $\mathbf{y}$ is the acquired k-space data, $\mathbf{F}_u$ is the undersampled Fourier transform, and η is a weighting parameter. The first term enforces the data consistency in the original space, and the second term is a classical pre-imaging formulation. Thus, (10) can be viewed as a regularized pre-imaging problem. Expanding the second term in (10), we arrive at the following problem

$$\begin{aligned}\min_{\mathbf{x}} \ & \frac{1}{2}\left\|\mathbf{y}-\mathbf{F}_u\mathbf{x}\right\|_2^2+\eta\sum_m k(\mathbf{x}_m,\mathbf{x}_m)\\ & -2\eta\sum_m\boldsymbol{\gamma}_m^T\mathbf{k}_{\mathbf{x}_m\mathbf{p}}+\eta\sum_m\boldsymbol{\gamma}_m^T\mathbf{K}_{\mathbf{p}}\boldsymbol{\gamma}_m,\end{aligned} \quad (11)$$

where $\boldsymbol{\gamma}_m=[\gamma_1^m \quad \gamma_2^m \quad \cdots \quad \gamma_L^m]^T$. Note that the last term in (11) is a constant since $\boldsymbol{\gamma}_m$ and $\mathbf{K}_{\mathbf{p}}$ are known.

C. Optimization Algorithm

Since most kernel functions are non-linear functions, the optimization problem (11) is difficult to solve. Here we focus on RBF kernel which leads to $k(\mathbf{x}_m, \mathbf{x}_m)=1$. Neglecting the constant terms in (11), we obtain the following minimization problem

$$\min_{\mathbf{x}} \ \frac{1}{2}\left\|\mathbf{y}-\mathbf{F}_u\mathbf{x}\right\|_2^2-2\eta\sum_m\boldsymbol{\gamma}_m^T\mathbf{k}_{\mathbf{x}_m\mathbf{p}}. \quad (12)$$

Introducing an auxiliary variable $\mathbf{d}=\mathbf{x}$, we arrive at the equivalent form of (12)

$$\begin{aligned}\min_{\mathbf{d},\mathbf{x}} \ & \frac{1}{2}\left\|\mathbf{y}-\mathbf{F}_u\mathbf{d}\right\|_2^2-2\eta\sum_m\boldsymbol{\gamma}_m^T\mathbf{k}_{\mathbf{x}_m\mathbf{p}}\\ \text{s.t.} \ & \mathbf{d}=\mathbf{x}.\end{aligned} \quad (13)$$

The scaled augmented Lagrangian function of (13) is

$$L_\rho(\mathbf{d},\mathbf{x},\mathbf{u})=\frac{1}{2}\left\|\mathbf{y}-\mathbf{F}_u\mathbf{d}\right\|_2^2-2\eta\sum_m\boldsymbol{\gamma}_m^T\mathbf{k}_{\mathbf{x}_m\mathbf{p}}+\frac{\rho}{2}\left\|\mathbf{d}-\mathbf{x}+\mathbf{u}\right\|_2^2, \quad (14)$$

where $\mathbf{u}$ is called the scaled dual variable of the Lagrangian multiplier and ρ is the penalty parameter. We derive the following iteration scheme:

$$\mathbf{d}^{k+1}=\arg\min_{\mathbf{d}}\frac{1}{2}\left\|\mathbf{y}-\mathbf{F}_u\mathbf{d}\right\|_2^2+\frac{\rho}{2}\left\|\mathbf{d}-\mathbf{x}^k+\mathbf{u}^k\right\|_2^2, \quad (15)$$

$$\mathbf{x}^{k+1}=\arg\min_{\mathbf{x}} \ \frac{\rho}{2}\left\|\mathbf{d}^{k+1}-\mathbf{x}+\mathbf{u}^k\right\|_2^2-2\eta\sum_m\boldsymbol{\gamma}_m^T\mathbf{k}_{\mathbf{x}_m\mathbf{p}}, \quad (16)$$

$$\mathbf{u}^{k+1}=\mathbf{u}^k+\mathbf{d}^{k+1}-\mathbf{x}^{k+1}. \quad (17)$$

The d-subproblem has a closed-form solution:

$$\mathbf{d}=\mathbf{F}^H(\mathbf{P}+\rho\mathbf{I})^{-1}\mathbf{F}\left[\mathbf{F}_u^H\mathbf{y}+\rho(\mathbf{x}-\mathbf{u})\right], \quad (18)$$

where $\mathbf{F}$ is the Fourier transform, $\mathbf{P}$ is the undersampling pattern in k-space, and $\mathbf{F}_u=\mathbf{PF}$.

The x-subproblem is still highly non-linear and we apply the fixed-point iteration scheme [22]. Note that the objective function in (16) is separable for different m. Taking the derivative of the objective function for each m, and performing basic algebra operations, we get the fixed-point iteration scheme

$$\mathbf{R}_m^T\mathbf{x}_m=\frac{4c\eta\sum_{l=1}^{L}\gamma_l^m k(\mathbf{R}_m^T\mathbf{x}_m,\mathbf{R}_m^T\mathbf{p}_l)\mathbf{R}_m^T\mathbf{p}_l+\rho\ \mathbf{R}_m^T\mathbf{R}_m(\mathbf{d}+\mathbf{u})}{4c\eta\left[\sum_{l=1}^{L}\gamma_l^m k(\mathbf{R}_m^T\mathbf{x}_m,\mathbf{R}_m^T\mathbf{p}_l)\right]+\rho\mathbf{I}}. \quad (19)$$

The final reconstruction is then obtained by

$$\mathbf{x}=\sum_m\mathbf{R}_m^T\mathbf{x}_m. \quad (20)$$

III. Simulation Results

We used a set of ASL perfusion data on calf muscle to evaluate the performances of the proposed method. The acquisition parameters are: TR/TE= 2.8/1.2ms, flip angle = 5°, FOV =160 × 112 mm^2, and matrix size = 112 × 100 × 20 (#FE×#PE×#frame). We applied the 1-D random down sampling pattern along the PE direction frame by frame. The central part of the k-space was fully sampled with a total of six PE lines for all frames. The net reduction factor was 3. Conventional CS reconstruction using PCA as the sparsifying transform was also performed for comparison. The images obtained from the full k-space data were used as the reference.

As stated previously, we used RBF kernel for KPCA and the RBF width parameter c was 20. KPCA needs a set of training signals. We randomly selected some temporal profiles from the images of previous iteration. In the initializing stage, the training signals were obtained from the low resolution images based on the fully sampled central k-space region. In this simulation, we set the number of training signals (L) to 1024, and the number of PCs (K) to 50. The regularization parameters were manually tuned to get the best results and we chose $\lambda = 0.001$ and $\rho = 0.01$.

Figure 1 shows the reconstruction results of the second frame, which is challenging for CS reconstruction due to the low signal to noise ratio (SNR). Compared with conventional CS reconstruction, the proposed method reduces the aliasing artifacts, especially in the regions indicated by arrows. The error images demonstrate that the proposed method produces more accurate reconstruction in most regions. We notice that the proposed method shows higher errors on the edges of structures than other regions. The reason is probably that the training signals are randomly selected and the number of training signals may not be enough to cover all temporal variations. Increasing the amount of training signals is expected to improve the reconstruction qualities, but it will raise the computational complexity as well. Another approach is to change the way of selecting training signals and make sure they can cover all kinds of temporal profiles.

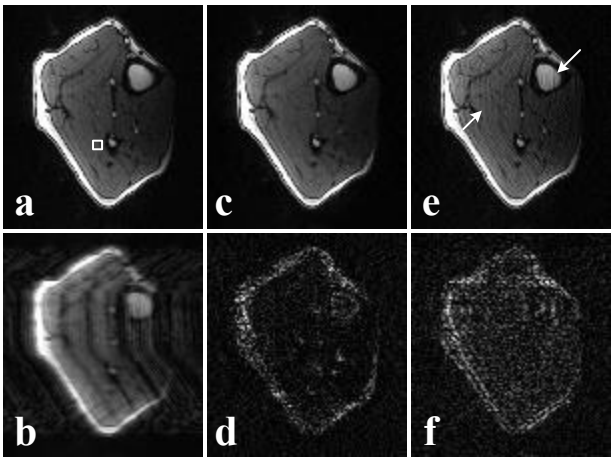

Figure 1. Comparison of reconstructions of the 2nd frame. (a) Reference image. (b) Result using zero-filling. (c) Result using the proposed method. (d) The error image of (c). (e) Result using conventional CS. (f) The error image of (e). For better visualization, the images are scaled by 2 and the errors are scaled by 10.

Figure 2 compares the average temporal curves of a selected region of interest (ROI) indicated in Figure 1(a). To get better visualization effects, we depict the zoom-in images of two parts. The result of the proposed method follows the reference curve closer than conventional CS method, suggesting that it is capable of capturing rapid kinetic information.

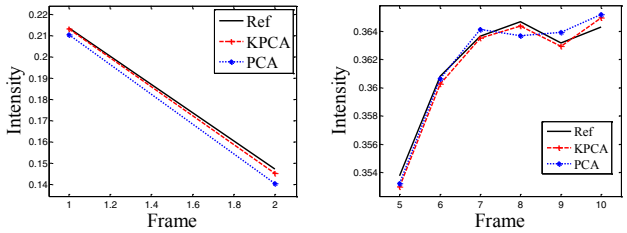

Figure 2. The average intensity of ROI v.s. frame curves. (a) Frame 1-2. (b) Frame 5-10.

IV. Conclusion

In this paper, we develop a novel undersampled dynamic MRI reconstruction model by integrating KPCA with the CS framework. We employ KPCA as the sparse representations of the dynamic variations in the feature space. The image sequence is then obtained by solving a modified pre-imaging problem. Retrospective simulation shows encouraging results that the proposed method can reduce the spatial aliasing artifacts and preserve the temporal variations. Future work will be on the selection of kernel functions and efficient optimization algorithms.

Acknowledgment

The authors would like to thank Dr. Jie Zheng from Washington University in St. Louis for providing the ASL perfusion muscle data.

References

[1] S. J. Riederer, T. Tasciyan, F. Farzaneh, J. N. Lee, R. C. Wright, and R. J. Herfkens, "MR fluoroscopy: Technical feasibility," *Magn. Reson. Med.*, vol. 8, pp. 1-15, 1988.

[2] Z.-P. Liang, and P. C. Lauterbur, "An efficient method for dynamic magnetic resonance imaging", *IEEE Trans. Med. Imag.*, vol. 13, pp. 677-686,1994.

[3] B. Madore, G. H. Glover, and N. J. Pelc, "Unaliasing by Fourier-encoding the overlaps using the temporal dimension (UNFOLD), applied to cardiac imaging and fMRI," *Magn. Reson. Med.*, vol. 42, pp. 813-828, 1999.

[4] J. A. d'Arcy, D. J. Collins, I. J. Rowland, et al., "Applications of sliding window reconstruction with Cartesian sampling for dynamic contrast enhanced MRI," *NMR Biomed.*, vol. 15, pp. 174-183, 2002.

[5] Z.-P. Liang, "Spatiotemporal imaging with partially separable functions," *in Proc. 2007 IEEE Int. Symp. Biomed. Imag*, pp. 988-991.

[6] J. Tsao, P. Boesiger, and K. P. Pruessmann, "k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations," *Magn. Reson. Med.*, vol. 50, pp. 1031-1042, 2003.

[7] J. Tsao and S. Kozerke, "MRI temporal acceleration techniques," *J. Magn. Reson. Imag.*, vol. 36, pp. 543-560, 2012.

[8] M. Lustig, J. M. Santos, D. L. Donoho, et al., "k-t SPARSE: High frame rate dynamic MRI exploiting spatiotemporal sparsity," *in Proc. 14th ISMRM*, pp. 2420.

[9] U. Gamper, P. Boesiger and S. Kozerke, "Compressed sensing in dynamic MRI," *Magn. Reson. Med.*, vol. 59, pp. 365-373, 2008.

[10] H. Jung, K. Sung, K. S. Nayak, et al., "k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI," *Magn. Reson. Med.*, vol. 61, pp. 103-116, 2009.

[11] M. Doneva, P. Börnert, H. Eggers, et al., "Compressed sensing reconstruction for magnetic resonance parameter mapping," *Magn. Reson. Med.*, vol. 64, pp. 1114-1120, 2010.

[12] D. Liang, E.V.R. Dibella, R.R. Chen, et al., "k-t ISD: Dynamic cardiac MR imaging using compressed sensing with iterative support detection," *Magn. Reson. Med.*, vol. 68, pp. 41-53, 2012.

[13] Y. Wang and L. Ying. "Compressed sensing dynamic cardiac cine MRI using learned spatiotemporal dictionary." *IEEE Trans. Biomed. Eng.*, vol. 61, pp. 1109-1120, 2014.

[14] S. Gao, I. W. Tsang, and L.-T. Chia, "Kernel sparse representation for image classification and face recognition," *in Proc. 2010 Euro. Conf. Comput. Vision,* pp. 1-14.

[15] Y. Chen, N. M. Nasrabadi, and T. D.Tran, "Hyperspectral image classification via kernel sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, pp. 217-231, 2013.

[16] Y. Zhou, K. Liu, R. E. Carrillo, K. E. Barner, and F. Kiamilev, "Kernel-based sparse representation for gesture recognition." *Pattern Recognition*, vol. 46, pp. 3208–3222, 2013.

[17] S. Gao, I. W. Tsang, and C. Liang-Tien, "Sparse representation with kernels," *IEEE Trans. Image Process.,* vol. 22, pp. 423-434, 2013.

[18] H. Qi and S. Hughes, "Using the kernel trick in compressive sensing: Accurate signal recovery from fewer measurements." *in Proc. 2011 IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 3940-3943.

[19] F.P. Anaraki and S. Hughes, "Kernel compressive sensing." *in Proc. 2011 IEEE Int. Conf. Image Process.*, pp. 494-498.

[20] Y. Zhou, Y. Wang and L. Ying, "A kernel-based compressed sensing approach to dynamic MRI from highly undersampled data." *in Proc.2013 IEEE Int. Symp. Biomed. Imag.*, pp. 310-313.

[21] B. Schlkopf and A. J. Smola, "Learning with kernels: support vector machines, regularization, optimization, and beyond." MIT Press, Boston, 2001.

[22] B. Scholkopf, A. Smola, and K.R. Muller, "Kernel principal component analysis," *in Proc. 1997 ICANN LNCS*, pp. 583-588.

[23] S. Mika, B. Schölkopf, A. Smola, K.L. Müller , M. Scholz, and G. Rätsch, "Kernel PCA and de-noising in feature spaces," *Adv. Neural. Inf. Process. Syst.*, vol. 11, pp. 536-542, 1999.