ACCELERATING MR PARAMETER MAPPING USING NONLINEAR MANIFOLD LEARNING AND SUPERVISED PRE-IMAGING

Yihang Zhou[1] Chao Shi[1], Fuquan Ren[2], Jingyuan Lyu[1], Dong Liang[3] and Leslie Ying[1, 4]

[1]Department of Electrical Engineering, State University of New York at Buffalo, Buffalo, NY, USA
[2]Department of Electrical and Electronic Engineering, Dalian University of Technology, Dalian, China
[3]Paul C. Lauterbur Research Center for Biomedical Imaging, Shenzhen Institute of Advanced Technology, Shenzhen, China
[4]Department of Biomedical Engineering, State University of New York at Buffalo, Buffalo, NY, USA

ABSTRACT

In this paper, we propose a new reconstruction framework that utilizes nonlinear models to sparsely represent the MR parameter-weighted image in a high dimensional feature space. Different from the prior work with nonlinear models where the image series is reconstructed simultaneously, each image at a specific time point is assumed to lie in a low-dimensional manifold and is reconstructed individually. The low-dimensional manifold is learned from the training images generated by the parametric model. To reconstruct each image, among infinite number of solutions that satisfy the data consistent constraint, the one that is closest to the manifold is selected as the desired solution. The underlying optimization problem is solved using kernel trick and split Bregman iteration algorithm. The proposed method was evaluated on a set of in-vivo brain T2 mapping data set and shown to be superior to the conventional compressed sensing methods.

Index Terms— MR parameter mapping, compressed sensing, nonlinear manifold learning, kernel PCA, regularized pre-image

1. INTRODUCTION

As a powerful quantitative imaging tool for tissue characterizations, MR parameter mapping has demonstrated great potential in various clinical applications [1]. However, to accurately obtain the MR parameters values, a large amount of contrast-weighted images need to be acquired which can lead to a relatively long acquisition time that limits the practical utility of MR parameter mapping.

A range of fast imaging techniques have been developed to accelerate the MR parameter mapping. These methods explicitly or implicitly use the parametric signal model and utilize various constraints associate with it. Typical examples include temporal smoothness constraint [3], sparsity or structured sparsity constraint [4]-[9], low rank constraint [10][11], contrast-weighting constraint [12]-[14], or a combination of them [15]-[17].

The above mentioned methods all use linear models for the constraints. In recent years, nonlinear models have been proposed in the compressed sensing framework and shown superiority on characterizing the nonlinearity in the signal models over linear models for dynamic imaging applications [17][19]. The core concept of nonlinear model is to map the original signal space to a high dimensional feature space via a nonlinear function and seek the corresponding sparse representation of the mapped signal in the feature space.

In this paper, we propose a novel framework to utilize nonlinear models to sparsely represent the unknown image. Different from the prior work with nonlinear models where the image series is reconstructed simultaneously, each image at a specific time point is assumed to lie in a low-dimensional manifold and is reconstructed individually. The low-dimensional manifold is learned from the training images generated by the parametric model. To reconstruct each image, among infinite number of solutions that satisfy the data consistent constraint, the one that is closest to the manifold is selected as the desired solution. The underlying optimization problem is solved using kernel trick [20] and split Bregman iteration algorithm [21]. Experimental results on T2 mapping using retrospectively undersampled data show that the proposed method is capable of recovering the T2 map at high reduction factors accurately when the conventional compressed sensing methods with linear models fail.

This paper is structured as follows. In Section 2, we describe in details how the proposed method decouples the reconstruction process into two sequential steps: kernel-based manifold learning and regularized pre-imaging with data consistency and sparsity constraint. Section 3 evaluates the proposed method using in-vivo T2-weighted brain datasets. The paper is concluded in Section 4.

2. PROPOSED METHOD

2.1 Problem Formulation

To describe the data acquisition process, we represent the MR parameter-weighted image at the *t-th* acquisition to be reconstructed as an $N \times 1$ vector $\mathbf{x}$. A finite number of measurements, denotes as $\mathbf{y}$, are collected whose *i-th* entry can be written as

$$y_i = <\mathbf{x}, \mathbf{f}_i> + \varepsilon_i \qquad (1)$$

where $<\cdot>$ denotes the inner product, $\{\mathbf{f}_i\}_{i=1}^n$ denotes i-th row of the undersampled Fourier measurement matrix $\mathbf{F}$, and ε_i denotes the complex white Gaussian noise. Since the measurement operation is under-determined, the direct recovery of $\mathbf{x}$ from measurement $\mathbf{y}$ is ill-posed.

To reveal the nonlinear sparseness of the parameter-weighted image $\mathbf{x}$, we define a nonlinear mapping $\Phi:\Re^p \to \mathfrak{I}^q$ that maps the original data space $\Re^p$ into a much higher, even infinite, dimensional feature space $\mathfrak{I}^q$ via a nonlinear function $\Phi(\cdot)$. We assume that $\mathbf{x}$ is nonlinearly d-sparse, meaning it can be sparsely represented by the first d largest coefficients in an orthonormal basis of the feature space $\mathfrak{I}^q$

$$\Phi(\mathbf{x}) \approx \sum_{k=1}^d \beta_k \mathbf{v}_k \qquad (2)$$

with β being the unknown coefficients, and $\{\mathbf{v}_k\}_{k=1}^q$ being the basis constructed using a set of training images $\{\mathbf{J}_i\}_{i=1}^L$ generated beforehand based on the parametric model via kernel PCA (kPCA) [22]. Specifically, in kPCA, each $\mathbf{v}_k \in \mathfrak{I}^q$ is represented in the expansion of

$$\mathbf{v}_k = \sum_{l=1}^L \alpha_l^k \Phi(\mathbf{J}_l) \qquad (3)$$

where $\{\alpha_i^k\}_{i=1}^L$ is the representation coefficients that can be obtained by solving the following eigenvalue problem

$$\mathbf{K}\alpha_k = \lambda \alpha_k \qquad (4)$$

where $\mathbf{K}$ is a $L \times L$ kernel matrix with the $(i,j)-th$ entry $\mathbf{K}_{i,j} = \kappa(\mathbf{J}_i,\mathbf{J}_j) = <\Phi(\mathbf{J}_i),\Phi(\mathbf{J}_j)>$ with $\kappa(\mathbf{J}_i,\mathbf{J}_j)$ being a kernel function and λ is the eigenvalues of $\mathbf{K}$. For simplicity, we assume that both $\Phi(\mathbf{J}_l)$ and $\mathbf{K}$ are centered in $\mathfrak{I}^q$. Otherwise we must replace $\Phi(\mathbf{J}_l)$ with the centered $\bar{\Phi}(\mathbf{J}_l) = \Phi(\mathbf{J}_l) - \sum_{l=1}^L \mathbf{J}_l/L$ and replace $\mathbf{K}$ with centered $\mathbf{K}^c = \mathbf{K} - \mathbf{1}_L\mathbf{K} - \mathbf{K}\mathbf{1}_L + \mathbf{1}_L\mathbf{K}\mathbf{1}_L$.

Based on the above assumption, the desired MR parameter-weighted image $\mathbf{x}$ lies in a low dimensional manifold where the key information of $\mathbf{x}$ is preserved by a small number of projection coefficients $\beta = [\beta_1,\beta_2,...,\beta_d]^T$. Therefore, we propose the following two step sequential reconstruction scheme to recover $\mathbf{x}$.

2.2 Nonlinear Manifold Learning

In this step, we aim to find the projection coefficients $\beta = [\beta_1,\beta_2,...,\beta_d]^T$ in the feature space $\mathfrak{I}^q$ by solving

$$\mathbf{s} = \mathbf{G}\beta + \delta. \qquad (5)$$

where $\mathbf{s}_i = \kappa(\mathbf{x},\mathbf{f}_i)$. Here we focus on a polynomial kernel

$$\kappa(\mathbf{x},\mathbf{f}_i) = (<\mathbf{x},\mathbf{f}_i>+\eta)^\gamma \approx (y_i+\eta)^\gamma \qquad (6)$$

where η is a positive scalar and γ is the polynomial order,

$\mathbf{G} = \mathbf{M\Lambda}$ with $\mathbf{M}$ an $n \times L$ matrix with the $(i,j)-th$ entries defined as $\mathbf{M}_{i,j} = \kappa(\mathbf{J}_j,\mathbf{f}_i)$ and $\mathbf{\Lambda}$ is an $L \times d$ matrix whose rows are formed by $\{\alpha_i^T\}_{i=1}^L$. The least square solution of Eq. (5) is given by

$$\beta = (\mathbf{G}^H\mathbf{G})^{-1}\mathbf{G}^H\mathbf{s}. \qquad (7)$$

Thereby we can obtain the unknown image in the feature space as $\mathbf{z} = \sum_{k=1}^d \beta_k v_k$.

2.3 Supervised Pre-Imaging

The objective of this step is to find the desired image $\mathbf{x}$ in the original space, which is a pre-image of $\mathbf{z}$ [23][24], that is, $\Phi(\mathbf{x}) = \mathbf{z}$. We also enforce that the pre-image satisfies the data consistency and transform sparse constraints. Therefore finding $\mathbf{x}$ is formulated as a supervised pre-image problem

$$\underset{\mathbf{x}}{\arg\min}\|\mathbf{Fx} - \mathbf{y}\|_2^2 + \lambda_1\|\mathbf{Hx}\|_1 + \lambda_2\|\Phi(\mathbf{x}) - \mathbf{z}\|_2^2, \qquad (8)$$

where $\mathbf{H}$ denotes the two-dimensional finite difference operation matrix. The first term represents the data consistency, the second term exploits the sparsity constraint, and the third term enforces the solution to lie on the manifold spanned by the predefined feature space.

Equation (8) can be highly nonlinear and nonconvex. However, when the polynomial kernel function has an odd order γ, the kernel function $k(\mathbf{x},\mathbf{f}_i) = f(<\mathbf{x},\mathbf{f}_i>)$ is invertible, and there is a unique solution to $\Phi(\mathbf{x}) = \mathbf{z}$:

$$\mathbf{x} = \sum_{i=1}^p f^{-1}\left(\sum_{l=1}^L c_l \kappa(\mathbf{J}_l,\mathbf{f}_i)\right)\mathbf{f}_i.$$

where $c_l = \sum_{k=1}^d \beta_k \alpha_l^k$. To solve the optimization problem, we rewrite Eq. (8) as

$$\underset{\mathbf{x}}{\arg\min}\|\mathbf{Fx} - \mathbf{y}\|_2^2 + \lambda_1\|\mathbf{U}\|_1 + \lambda_2\left\|\mathbf{V} - \sum_{i=1}^p f^{-1}\left(\sum_{l=1}^L c_l\kappa(\mathbf{J}_l,\mathbf{f}_i)\right)\mathbf{f}_i\right\|_F^2$$

$$\text{subject to } \mathbf{V} = \mathbf{x}, \text{ and } \mathbf{U} = \mathbf{HV}, \qquad (9)$$

where $\mathbf{V}$ and $\mathbf{U}$ are called the auxiliary variables. Eq. (9) now can be solved iteratively by minimizing the cost function with respect to $\mathbf{x}$, $\mathbf{V}$ and $\mathbf{U}$ separately using split Bregman algorithm without venturing into feature space [23]. Specifically, at the μ-th iteration, the scaled augmented Lagrangian function of Eq. (9) is

$$L_\vartheta(\mathbf{x},\mathbf{V},\mathbf{U}) = \frac{1}{2}\|\mathbf{Fx} - \mathbf{y}\|_2^2 + \frac{\vartheta_1}{2}\left\|\mathbf{V} - \sum_{i=1}^p f^{-1}\left(\sum_{l=1}^L c_l\kappa(\mathbf{J}_l,\mathbf{f}_i)\right)\mathbf{f}_i\right\|_F^2$$

$$+ \frac{\vartheta_2}{2}\|\mathbf{x} - \mathbf{V} - \mathbf{D}_1^\mu\|_2^2 + \frac{\vartheta_2}{2}\|\mathbf{U} - \mathbf{HV} - \mathbf{D}_2^\mu\|_2^2 + \vartheta_3\|\mathbf{U}\|_1 \qquad (10)$$

with

$$\begin{aligned}\mathbf{D}_1^{(\mu+1)} &= \mathbf{D}_1^\mu + (\mathbf{V}^{(\mu+1)} - \mathbf{x}^{(\mu+1)})\\ \mathbf{D}_2^{(\mu+1)} &= \mathbf{D}_2^\mu + (\mathbf{HV}^{(\mu+1)} - \mathbf{U}^{(\mu+1)})\end{aligned} \qquad (11)$$

We now describe the three steps to be performed to solve Eq. (10).

- ***x-step: Optimization with respect to x***

In this step, Eq. (10) reduces to

$$\mathbf{x}^{(\mu+1)} = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{Fx} - \mathbf{y}\|_2^2 + \frac{\vartheta_2}{2}\|\mathbf{x} - \boldsymbol{\zeta}^{(\mu)}\|_2^2 \qquad (12)$$

where $\boldsymbol{\zeta}^{(\mu)} = \mathbf{V}^{(\mu)} + \mathbf{D}_1^{(\mu)}$. Eq. (12) has a closed-form solution

$$\mathbf{x}^{(\mu+1)} = \mathbf{F}^H(\mathbf{P} + \vartheta_2\mathbf{I})^{-1}\mathbf{F}\left[\mathbf{F}_u^H\mathbf{y} + \vartheta_2\boldsymbol{\zeta}^{(\mu)}\right]. \qquad (13)$$

- **V-step: Optimization with respect to V**

In this step, Eq. (16) reduces to

$$\begin{aligned}\mathbf{V}^{(\mu+1)} = \arg\min_{\mathbf{V}} &\frac{\vartheta_2}{2}\|\mathbf{x}^{(\mu+1)} - \mathbf{V} - \mathbf{D}_1^{(\mu)}\|_2^2 + \frac{\vartheta_2}{2}\|\mathbf{U}^{(\mu)} - \mathbf{HV} - \mathbf{D}_2^{(\mu)}\|_2^2 \\ &+ \frac{\vartheta_1}{2}\left\|\mathbf{V} - \sum_{i=1}^{p} f^{-1}\left(\sum_{l=1}^{L} c_l\kappa(\mathbf{J}_l,\mathbf{f}_i)\right)\mathbf{f}_i\right\|_F^2\end{aligned} \qquad (14)$$

Taking the first order derivative of Eq. (14) with respect to **V** yields

$$(\mathbf{x}^{(\mu+1)} - \mathbf{V} - \mathbf{D}_1^{(\mu)}) + \mathbf{H}^T(\mathbf{U}^{(\mu)} - \mathbf{HV} - \mathbf{D}_2^{(\mu)}) + \frac{\vartheta_1}{\vartheta_2}(\mathbf{V} - \mathbf{Q}) = 0 \quad (15)$$

where $\mathbf{Q} = \sum_{i=1}^{p} f^{-1}\left(\sum_{l=1}^{L} c_l\kappa(\mathbf{J}_l,\mathbf{f}_i)\right)\mathbf{f}_i$. The solution to Eq. (15) is given by

$$\mathbf{V}^{(\mu+1)} = \left[\mathbf{x}^{(\mu+1)} - \mathbf{D}_1^{(\mu)} + \mathbf{H}^T(\mathbf{U}^{(\mu)} - \mathbf{D}_2^{(\mu)}) - \frac{\vartheta_1}{\vartheta_2}\mathbf{Q}\right]\left[(1 + \frac{\vartheta_1}{\vartheta_2})\mathbf{I} + \mathbf{H}^T\mathbf{H}\right]^{-1}. \quad (16)$$

- **U-step: Optimization with respect to U**

The last step we have to consider is

$$\mathbf{U}^{(\mu+1)} = \arg\min_{\mathbf{U}} \vartheta_3\|\mathbf{U}\|_1 + \frac{\vartheta_2}{2}\|\mathbf{U} - \mathbf{HV}^{(\mu+1)} - \mathbf{D}_2^{(\mu)}\|_2. \quad (17)$$

This is a typical L1 minimization problem which can be easily solved by any off-the-shelf optimization method.

Finally, by iterating through the abovementioned three steps we may accurately recover the desired MR parameter weighted image **x** from the *t-th* acquisition.

3. EXPERIMENTS AND RESULTS

We evaluate the proposed method using T2 mapping experiments, where the signal is described by an exponential decay. However, the method can be generalized for any other parameter mapping models. A set of brain data was acquired on a 3T scanner using a 12-channel head coil with a turbo spin echo sequence (matrix size = 192×192, FOV = 192×192 mm, slice thickness = 3.0 mm, ETL = 16, ΔTE = 8.8 ms, TR = 4000 ms, bandwidth = 362 Hz/pixel). The k-space obtained at the first echo time was not used due to its hypointense. To simulate the reduced acquisition retrospectively, the k-space data was randomly undersampled along the phase encoding direction at each echo time with net reduction factors of 4 and 8, respectively. Different sampling patterns were used at different TEs. The T2 map obtained by pixel-wise fitting using the Levenberg-Marquardt algorithm from the fully sampled data was used

as the gold standard for comparison between the proposed method and the conventional compressed sensing method with principle component analysis as the sparsifying transform (CS-PCA).

As previously stated, the proposed method requires training images to estimate the basis function in the feature space. The k-space at the second echo time was fully sampled and used to form an image $\boldsymbol{\rho}$. Similar to MR fingerprinting [25], the training images were generated by $\boldsymbol{\rho}e^{(-TE/T2)}$ with a number of possible T2 map based on the segmentation of different tissues. A total of 976 training images were generated for each echo time. We used the polynomial kernel with $\eta = 15$ and $\gamma = 5$. The regularization parameters were manually tuned to get the best result and we chose $\vartheta_1 = 0.5$, $\vartheta_2 = 0.2$ and $\vartheta_3 = 0.001$.

Fig. 1 shows the reconstructed T2 maps and their corresponding pixel-wise error maps. The overall percentage errors are shown on the left top corner of the error maps. For better visualization, the error maps are magnified by 6. Visually the T2 maps obtained from the proposed method are indistinguishable from those obtained from the fully sampled data while the ones obtained from conventional CS reconstruction using linear PCA shows significant aliasing artifacts, especially in the white matter. We noticed that for the proposed method relatively higher error appears on the gray matter. This is likely due to the limited number of training images. Increasing the number of training images is expected to improve the reconstruction quality but at a cost of increased computational complexity. Another possibility is to improve the tissue segmentation so that the training images can be closer to the unknown image. To quantitatively assess our proposed method, the overall percentage errors were shown on the left top corner of the error maps. The results from the proposed method show good agreement with those from the fully sampled reference and are consistent for both undersampling factors of 4 and 8.

4. CONCLUSION

In this paper, a novel reconstruction method is proposed to accelerate the MR quantitative imaging. Compared to the existing compressed sensing methods, the proposed method constrains both the linear and nonlinear sparsity of the images using nonlinear manifold learning and supervised pre-imaging. The underlying optimization problem is solved by kernel trick and split Bregman algorithm. Experimental results show that the proposed method can accurately recover MR parameter maps from highly undersampled datasets. Future work will investigate selection of kernel functions and evaluation on other MR parametric models.

5. ACKNOWLEDGMENT
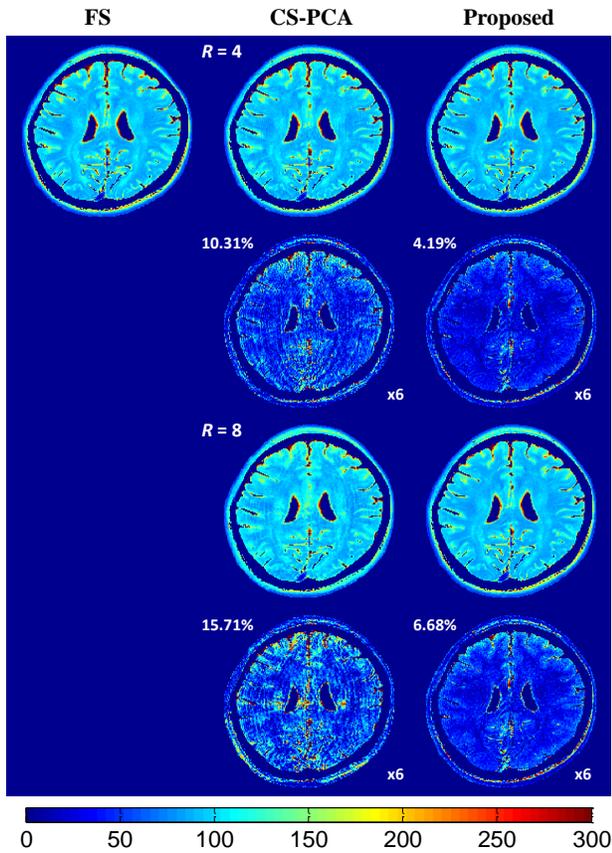
Figure 1. Estimated T2 maps and pixel-wise error maps of the brain dataset using the reconstruction from fully sampled data, CS-PCA and the proposed method with reduction factors of 4 and 8.

6. REFERENCES

[1] H.-L. Margaret Cheng, N. Stikov, N. R. Ghugre, and G. A. Wright, "Practical medical applications of quantitative MR relaxometry," *J. Magn. Reson. Imag.,* 2012; 36: pp. 805–824, 2012.

[2] M. Lustig, D.L. Donoho, and J.M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Mag. Reson. Med. 2007,* 58: pp. 1182–1195.

[3] J.V. Velikina, A. L. Alexander, A. Samsonov, "Accelerating MR parameter mapping using sparsity-promoting regularization in parametric dimension," *Magn. Reson. Med.* 2013; 70:1263–1273.

[4] M. Doneva, P. Bornert, H. Eggers, C. Stehning, J. Senegas, A. Mertins, "Compressed sensing reconstruction for magnetic resonance parameter mapping," *Magn. Reson. Med.* 2010; 64:1114–1120.

[5] L. Feng, R. Otazo, H. Jung, J. H. Jensen, J.C. Ye, D. K. Sodickson, D. Kim, "Accelerated cardiac T2 mapping using breath-hold multiecho fast spin-echo pulse sequence with k-t FOCUSS," *Magn. Reson. Med.* 2011; 65:1661–1669.

[6] Bilgic B, Goyal VK, Adalsteinsson E. "Multi-contrast reconstruction with Bayesian compressed sensing," *Magn. Reson. Med.* 2011; 66:1601–1615.

[7] A. Majumdar, R. K. Ward, "Accelerating multi-echo T2 weighted MR imaging: analysis prior group-sparse optimization," *J. Magn. Reson. Imaging.* 2011; 210:90–97.

[8] J. Yuan, D. Liang, F. Zhao, Y. Li, Y.-X. Xiang, L. Ying, "k-t ISD compressed sensing reconstruction for T1 mapping: a study in rat brains at 3T," *in Proc. ISMRM 2012.* p. 4197.

[9] J. Huang, C. Chen, L. Axel, "Fast multi-contrast MRI reconstruction," *in Proc.* 15th *MICCAI* 2012. pp. 281–288.

[10] F. H. Petzschner, I. P. Ponce, M. Blaimer, P. M. Jakob, F. A. Breuer. Fast MR parameter mapping using k-t principal component analysis. *Magn. Reson. Med.* 2011; 66:706–716.

[11] T. Zhang, J. M. Pauly, I. R. Levesque, "Accelerating parameter mapping with a locally low rank constraint," in press, *Magn. Reson. Med.* 2014.

[12] J. P. Haldar, D. Hernando, Z.-P. Liang, "Super-resolution reconstruction of MR image sequences with contrast modeling," *In Proc. ISBI*, 2009. pp. 266–269.

[13] K. Block, M. Uecker, J. Frahm, "Model-based iterative reconstruction for radial fast spin-echo MRI," *IEEE Trans. Med. Imaging* 2009; 28:1759–1769.

[14] T. J. Sumpf, M. Uecker, S. Boretius, J. Frahm, "Model-based nonlinear inverse reconstruction for T2 mapping using highly undersampled spin-echo MRI," *J. Magn. Reson. Imaging.* 2011; 34:420–428.

[15] C. Huang, C. G. Graff, E. W. Clarkson, A. Bilgin, M. I. Altbach, "T2 mapping from highly undersampled data by reconstruction of principal component coefficient maps using compressed sensing," *Magn. Reson. Med.* 2012; 67: 1355–1366.

[16] B. Zhao, W. Lu, Z.-P. Liang, "Accelerated parameter mapping with low rank and sparsity constraints," *in press, Magn. Reson. Med.2014.*

[17] B. Zhao, F. Lam, Z.-P. Liang, "Model-based MR parameter mapping with sparsity constraints: parameter estimation and performance bounds," in press, *IEEE Trans. Med. Imaging* 2014.

[18] Y. Zhou, Y. Wang, L. Ying, "A kernel-based compressed sensing approach to dynamic MRI from highly undersampled data," *In Proc. ISBI*, 2013, pp. 310-313.

[19] Y. Wang, L. Ying, "Undersampled Dynamic Magnetic Resonance Imaging using Kernel Principal Component Analysis," *In Proc. IEEE EMBC,* 2014 pp.1533-1536.

[20] 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.

[21] N. Nguyen, J. Chen, C. Richard, P. Honeine, and C. Theys. "Supervised nonlinear unmixing of hyperspectral images using a pre-image methods," *EAS Publications Series* 2013, 59, 417-437.

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

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

[24] 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.

[25] D. Ma, V. Gulani, N. Seiberlich, K. Liu, J. L. Sunshine, J. L. Duerk, and M. A. Griswold. Magnetic resonance fingerprinting. *Nature* 2013, 495(7440): pp. 187-192.