Multi-Kernel Regression Imputation on Manifolds via Bi-linear Modeling

Gaurav Shetty[1], **Konstantinos Slavakis**[1], Loris Cannelli[2], Ukash Nakarmi[3], Gesualdo Scutari[4], Leslie Ying[1,5]

UB: [1]**Dept. of Electrical Eng.**, [5]Dept. of Biomedical Eng., [2]Dalle Molle Institute for Artificial Intelligence, [3]Stanford University: Dept. of Electrical Eng. & Radiology; [4]Purdue University: Dept. of Industrial Eng.

University at Buffalo
Department of Electrical Engineering

2nd Buffalo Wireless Day 2020

Introduction

What is missing data problem?

- Missing Data - Gaps in the observed/acquired data or data loss due to data delivery.
- Data acquisition and delivery processes are affected by sensor faults, connection errors in sensor networks, physical acquisition constraints, security attacks, environmental factors, and so on ...

Applications:

- Imaging Applications : MRI, CT, Remote Sensing, Seismic Imaging.
- Internet of Things (IoT) : Data fusion in healthcare monitoring, Internet of Vehicles (IoV).
- Social Media : Recommender Systems, Sentiment Analysis, Social Network Analysis.

Impact:

- Imaging Applications : Incomplete data scanning leads to distorted, artifact-induced, low-resolution medical or geological imaging, not acceptable; Expensive and time consuming data acquisition schemes.
- Internet of Things : Can adversely affect the protocol implementation which plays a crucial role in remote, energy dependent sensor networks.
- Social Media : Loss of correlation information between features; Can lead to learning of incorrect models for data analysis which can further lead to incorrect biases and interpretations.

Main Contributions

1. This study proposes a low-rank kernel scheme for reconstructing data on manifolds which comes in handy to impute the missing values encountered in data acquisition processes.
2. This scheme is novel for its incorporation of kernel functions designed not only for real valued data but also complex valued data acquired commonly in imaging applications.
3. This scheme proposes a data model based on classical kernel arguments and employs a bi-linear model which avoids the need for a pre-imaging process and the need for choosing an optimal kernel function which can vary in accordance to data.
4. The framework doesn't rely on the availability of a fully-observed training data (unlike deep learning schemes) rather uses the partially observed data itself to learn a data model from which the missing values can be filled.
5. The data model can be applied to various healthcare, social media and IoT applications.
6. The efficacy of the study is validated for the MRI application where good quality medical images are generated from partially observed scanner data.
7. This study outperforms other state-of-the-art reconstruction schemes for the dMRI recovery problem.

Main Idea

- Consider $\mathbf{y}_t$ = Measured time series data at time t.
- Non-linearly map all $\mathbf{y}_t$ to feature space where the similarity between the data points is exploited via kernel function $\kappa(\cdot,\cdot)$.
- Based on the assumption feature maps lie on the smooth low-dimensional manifold learn the compressed latent representations for the same.
- Latent Representations - Use the concept of tangent space to locally and affinely combine the neighboring points to describe each point on the manifold.
- Compression imposes low-rank structure to the data model, highly desirable in reconstruction problems.
- Reconstruct the high dimensional data point from latent representations.

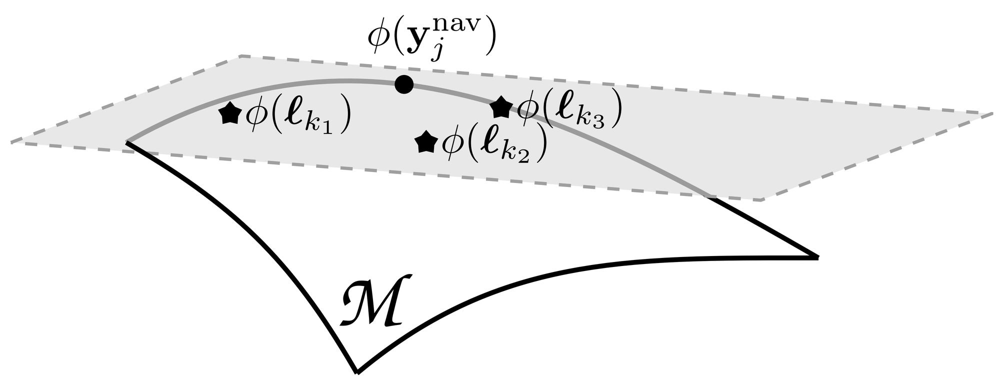


Figure: Low-dimensional Manifold embedded in the feature space on which the feature maps lie

Mathematical Formulation: The Dynamic-MRI Case

- MR scanner time series data: $\mathbf{Y}$.
- dMRI Image series: $\mathbf{X} = \mathcal{F}^{-1}(\mathbf{Y})$ related by inverse 2D Fourier Transform.
- Partially observed $\mathcal{S}(\mathbf{Y}) \implies$ Distorted, aliased, artifact induced $\mathbf{X}$.
- $\mathcal{F}_t(\cdot)$: 1D Fourier Transform along temporal axis.

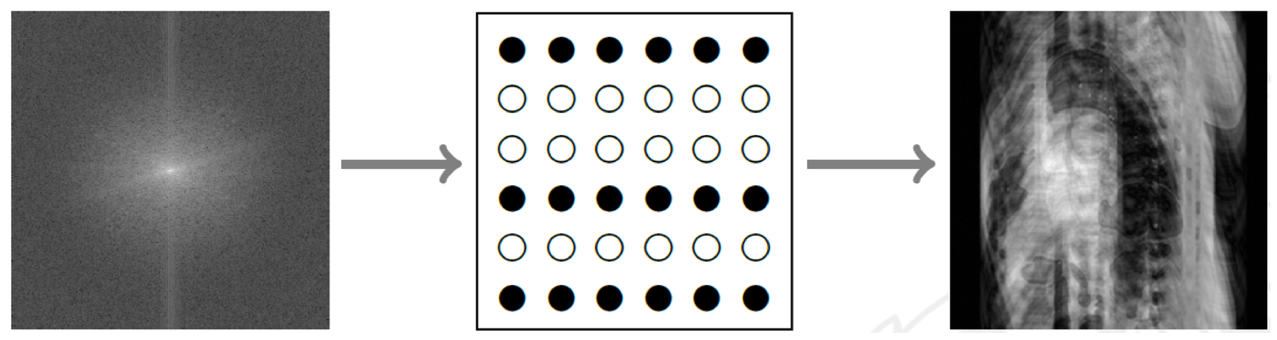


Figure: Visual Representation of the data acquisition process

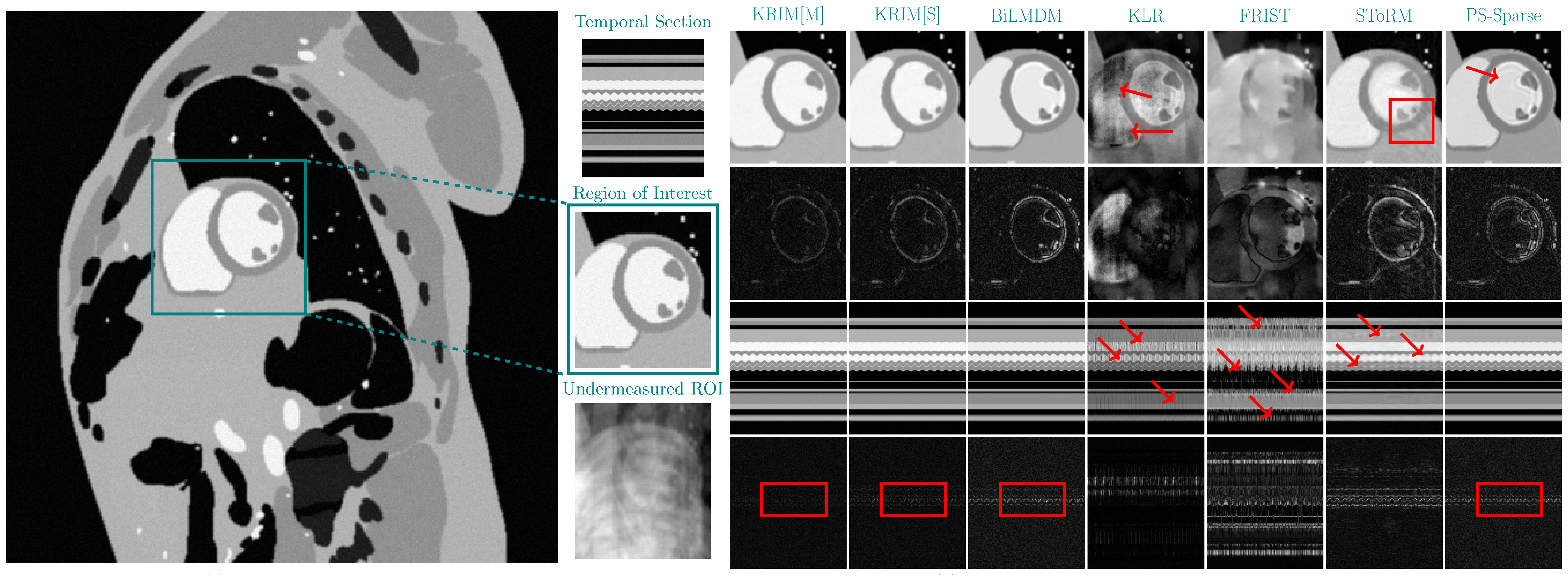
- Kernel Dictionary $\{\mathbf{K}\}_{i=1}^{M}$ generated from heavily sampled area of the sensor data (Gaussian & Polynomial).
- Compressed Kernel Dictionary $\{\hat{\mathbf{K}}\}_{i=1}^{M}$ using robust sparse embedding in the feature space.
- Kernel Dictionary: Formulation allows for a data adaptive kernel expressed as an affine combination of a wide range of kernels (**no need to tune for optimal kernels**)
- **Modeling:** $\mathbf{X} = \sum_{i=1}^{M} \mathbf{D}_i \hat{\mathbf{K}}_i \mathbf{B}_i$; Shorthand Notation: $\mathbf{X} = \tilde{\mathbf{D}}\tilde{\mathbf{K}}\tilde{\mathbf{B}}$

Inverse Problem Formulation

$$\min_{\mathbf{X},\tilde{\mathbf{D}},\tilde{\mathbf{B}},\mathbf{Z}} \overbrace{\tfrac{1}{2}\|\mathbf{X} - \tilde{\mathbf{D}}\tilde{\mathbf{K}}\tilde{\mathbf{B}}\|_{\mathrm{F}}^2}^{\text{Data-fit Term}} + \overbrace{\lambda_1\|\tilde{\mathbf{B}}\|_1}^{\substack{\text{Sparsity on}\\ \text{representations}}} + \overbrace{\tfrac{\lambda_2}{2}\|\mathbf{Z} - \mathcal{F}_{\mathrm{t}}(\mathbf{X})\|_{\mathrm{F}}^2 + \lambda_3\|\mathbf{Z}\|_1}^{\substack{\text{Sparsity along temporal}\\ \text{direction due to periodicity}}}$$

$$\text{s.to } \underbrace{\|\tilde{\mathbf{D}}\mathbf{e}_i\| \leq C_{\mathrm{D}}\ \forall i \in \{1,\ldots,Md\}}_{\text{Bounding Constraint}};\ \underbrace{\mathbf{1}_{N_\ell}^{\top}\mathbf{B}_i = \mathbf{1}_{N_{\mathrm{fr}}}\ \forall i \in \{1,\ldots,M\}}_{\text{Affine Constraints}};\ \underbrace{\mathcal{S}(\mathbf{Y}) = \mathcal{F}\mathcal{S}(\mathbf{X})}_{\text{MRI Data Relation}}$$

Numerical Results



(a) Ground Truth and Measured Data (b) Comparisons across state-of-the-art techniques

Figure: The efficacy of the proposed scheme (KRIM), in the multi-kernel [M] and single kernel [S] setting is validated against various state-of-the-art techniques. The reconstructions are achieved for the MRXCAT phantom simulating and acceleration rate of 8x. The red markings highlight the distortions in reconstructions of the competitive methods.

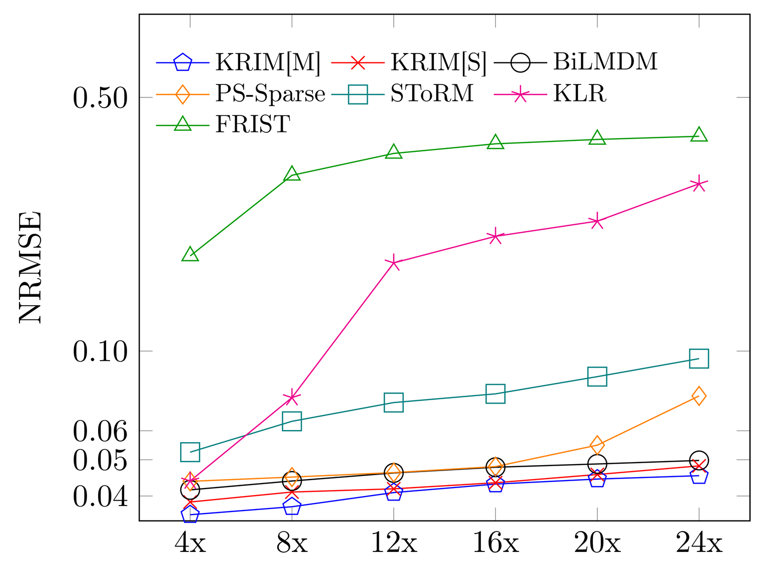


Figure: NRMSE values vs. acceleration rates

	NRMSE	SSIM	HFEN	M1	M2
KRIM[M]	**0.0372**	**0.9052**	**0.1058**	**37.95**	$\mathbf{1.5 \times 10^6}$
KRIM[S]	0.0407	0.8869	0.1198	37.65	1.4×10^6
BiLMDM	0.0438	0.8697	0.1314	28.69	1.1×10^6
KLR	0.0744	0.6704	0.1580	34.98	1.4×10^6
FRIST	0.3055	0.4572	0.7006	13.50	7.7×10^5
SToRM	0.0644	0.7743	0.2478	31.06	$\mathbf{1.5 \times 10^6}$
PS-Sparse	0.0449	0.8864	0.1338	30.66	1.2×10^6

Table: Quantitative Performance Analysis for MRXCAT Phantom (Acceleration Rate: 8x)

Validation

- Validation Numeric Metrics: NRMSE (voxel reconstruction error), HFEN (edge reconstruction error), M1 & M2 (sharpness measure) and SSIM (structural similarity).
- Quantitatively, the proposed kernel schemes present the best numbers when compared to the state-of-the-art methods.
- Qualitatively, the proposed scheme produces image reconstructions which are high-resolution, distortion-free, aliasing-free, artifact free and are very similar to the gold standard in regards to contrast and image structure.
- The proposed scheme consistently outperforms the other schemes for increasing number of missing values in the scanner data (acceleration rate).

Related Work

- G. Shetty, *etal.*, Bi-Linear Modeling of Data Manifolds for Dynamic-MRI Recovery. IEEE TMI, March 2020.
- G. Shetty, *etal.*, Kernel Bi-Linear Modeling on Data Manifolds: Dynamic-MRI Recovery. EUSIPCO 2020, Jan 2021.

*This research was supported in part by the NSF grants 1514056 and 1718796.