Sparsity Methods in Undersampled Dynamic Magnetic Resonance Imaging

Simon Arridge[1]
Joint work with: Benjamin Trémoulhéac [2]

[1]Department of Computer Science, University College London, UK
[2]Department of Medical Physics, University College London, UK

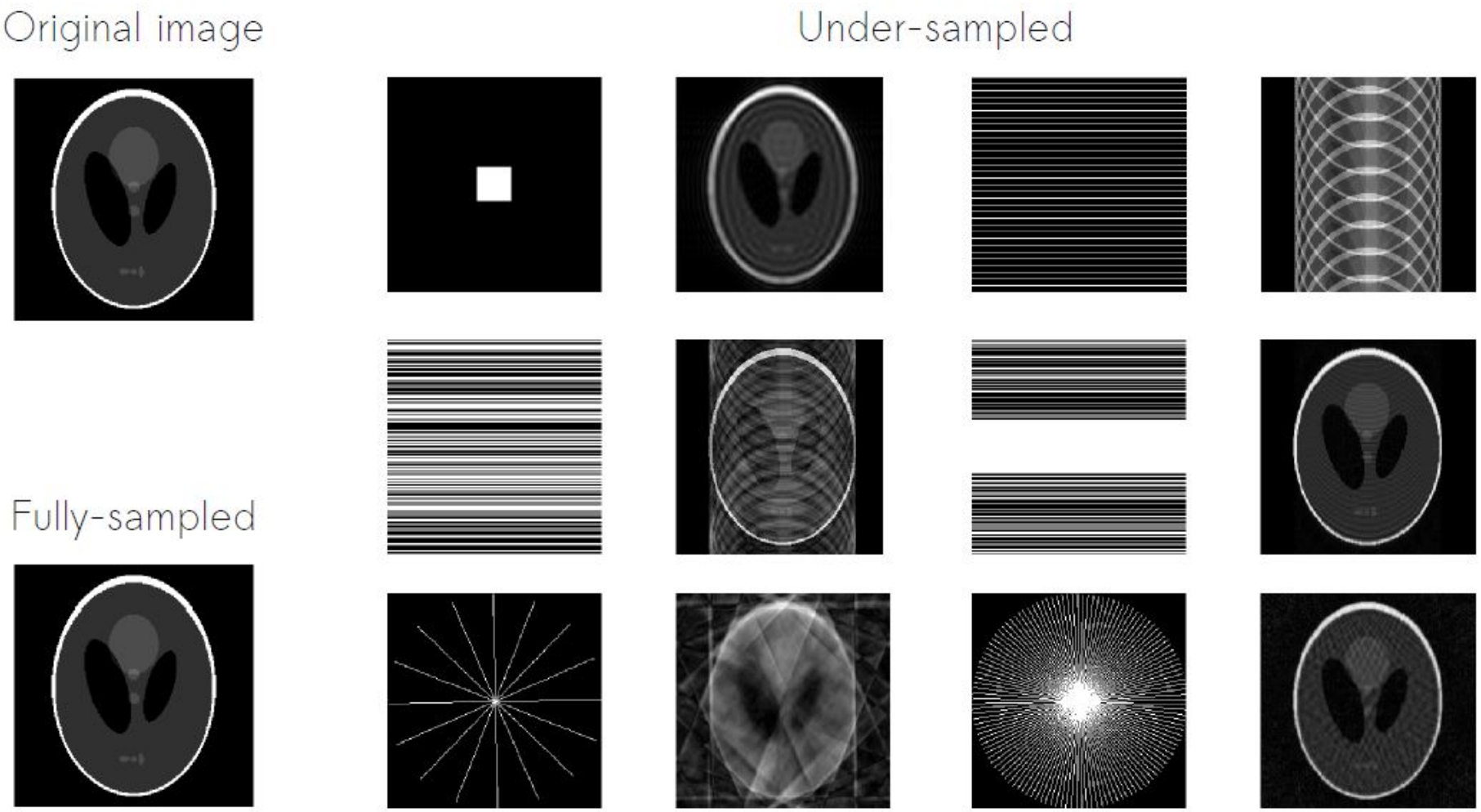
Sparse Tomography Seminar, March 26th, 2014

Outline

Outline

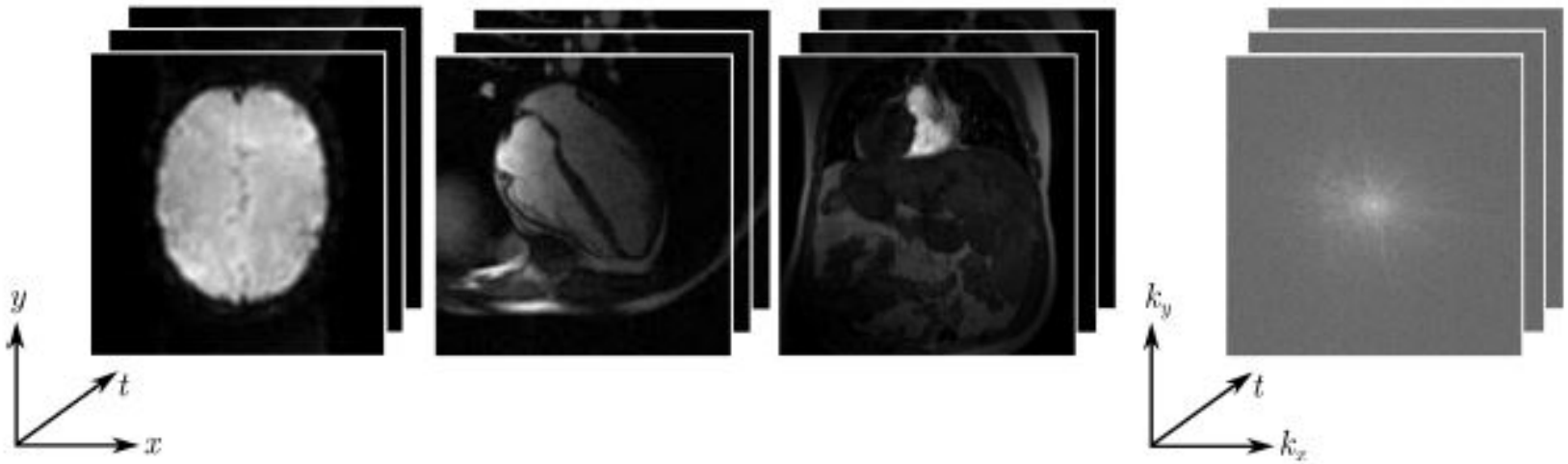
Introduction

Fully vs Under Sampled MRI



Introduction

Dynamic MRI from partial measurements



The imaging equation in dynamic MRI can be written as,

$$S(\boldsymbol{k}, t) = \int \gamma(\boldsymbol{x}, t) \mathrm{e}^{-\mathrm{i}2\pi(\boldsymbol{k}\cdot\boldsymbol{x})} \mathrm{d}\boldsymbol{x} + n(\boldsymbol{k}, t)$$

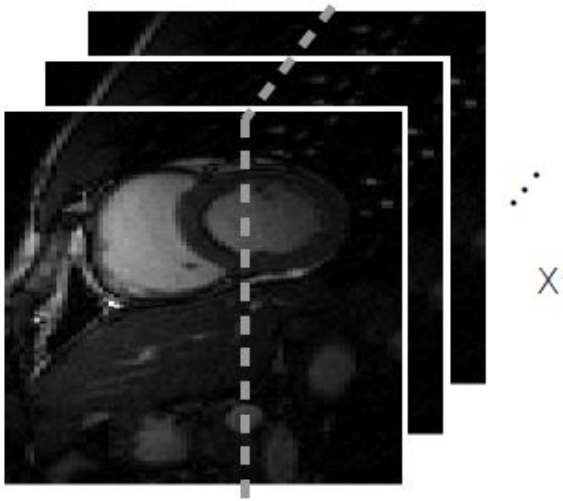
We assume the noise can be modelled by an additive white Gaussian distribution on both real and imaginary components (with i.i.d. random variables) .

Reconstructing $\gamma(\boldsymbol{x}, t)$ from a limited number of measurements of $S(\boldsymbol{k}, t)$ (*sub-Nyquist data*) is the inverse problem of interest.

Introduction

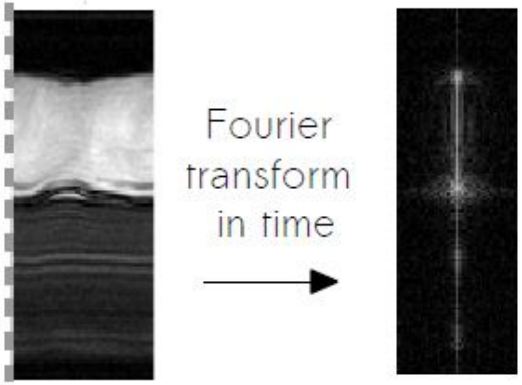
Compressed Sensing in Dynamic MRI

Spatiotemporal signal $\gamma(\mathbf{x}, t)$



x-y-t domain

Sparsity

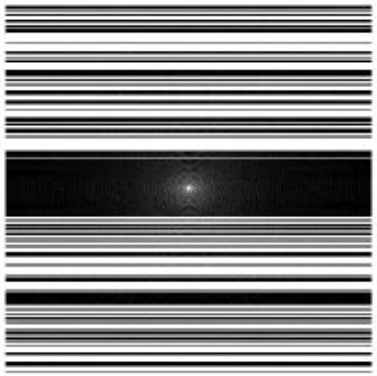


Fourier transform in time

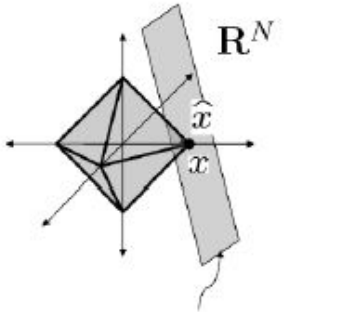
x-t domain

x-f domain

Random measurements



L1 reconstruction



$\{x' : \ y = \Phi x'\}$

Lustig et al, ISMRM 2006
Gamper et al, MRM 2008
Many others...

Introduction

Fast imaging methods previously proposed

Many techniques to tackle this inverse problem in dynamic MRI rely on the assumption that a Fourier transform along the temporal dimension (often called $(\boldsymbol{x} - f)$-space) returns an approximately sparse signal, because the original images in time exhibit significant correlation and/or periodicity.

This prior knowledge about a sparse $(\boldsymbol{x} - f)$-space has been used in techniques such as

- UNFOLD : (Madore *et.al.*, 1999) uses lattice under-sampling scheme that makes aliasing artefacts in x-f domain easily removable with a simple filter
- $\boldsymbol{k} - t$-BLAST : (J.Tsao, *et.al*, 2003), uses variable-density sampling scheme that provide lowspatial resolution approximation of the x-f space (training data) which gives a rough estimate of the signal distribution, and is then used to guide the reconstruction

Additionally the compactness of the signal distribution is implicitly exploited.

Compressed sensing suggests that if the signal of interest is sparse (in some domain or in its own), it is possible under some assumptions to reconstruct the signal exactly with high probability with many fewer samples than the standard Shannon-Nyquist theory recommends.
CS has been applied to MRI (M. Lustig, D. L. Donoho, and J. M. Pauly 2007) and in particular techniques have been developed specifically for dynamic MRI, such as

- $\boldsymbol{k} - t$-SPARSE : (M. Lustig, *et.al.* 2003,2006), use random under-sampling scheme to produce incoherent, noisy like artefacts that are removed by denoising via sparsity (l1 norm)
- $\boldsymbol{k} - t$-FOCUSS : (H. Jung, *et.al.*, 2007,2009), first estimates a low-resolution version of the $\boldsymbol{x} - f$-space prior to a CS reconstruction using the FOCUSS algorithm (I. F. Gorodnitsky and B. D. Rao, 1997), a general estimation method.

A finite-dimensional spatio-temporal MRI model is denoted

$$y = E(\boldsymbol{x}) + \boldsymbol{n}$$

where

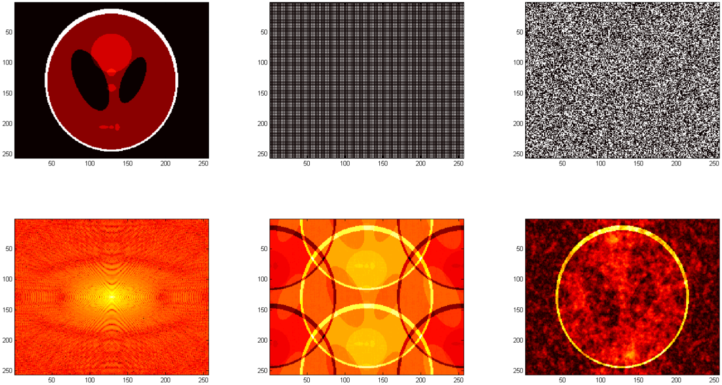
- $y \in \mathbb{C}^p$ represents the stacked (k-t)-space measurements vector,
- $E : \mathbb{C}^{N_x N_y N_t} \rightarrow \mathbb{C}^p$ is the MRI encoding operator modelling both the sub-Nyquist sampling and Fourier transform with $P \ll N_x N_y N_t$,
- $\in \mathbb{C}^{N_x N_y N_t}$ is the vectorized dynamic sequence (N_t images of dimensions $N_x \times Ny$) to recover, and
- $\boldsymbol{n} \in \mathbb{C}^p$ is the noise vector.

The task of finding $\boldsymbol{x}$ is a discrete linear illposed inverse problem.

Outline

Antialiasing

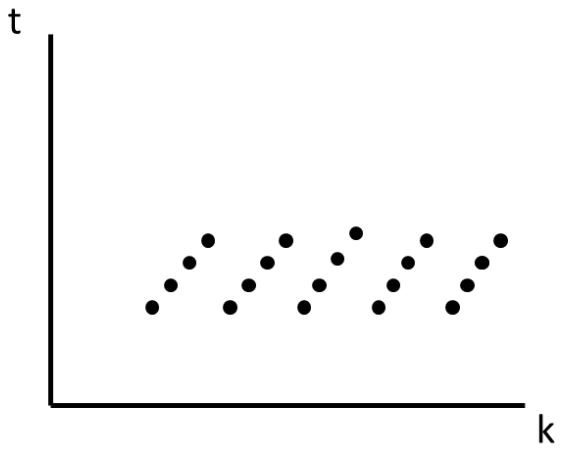
- Regular subsampling leads to aliasing
- Random sampling offsets aliasing to noise



KT-BLAST

KT-BLAST (Broad-use Linear Acquisition speed-up Technique) : J.Tsao, P.Boesinger, K.P. Preuessman, *Mag. Res. Med*, 2003

- **k**-space sampled regularly, interleaving samples in time
- ignoring time implies a high-resolution image
- low resolution used as prior to “unalias” the data.



In $\mathbf{x} - f$-space, aliasing requires solution of an undertermined problem

$$\rho^{\text{alias}} = A\rho$$

1. Put $\rho_{f=0}$ as the "DC" image (integral over time).
2. Estimate a prior distribution $M = \text{diag}\left[\rho^{\text{lowres}}\right]$
3. Solve $\rho^{\text{alias}} = AM(M^{-1}\rho)$
   $\Rightarrow \quad \rho = M^2 A^{\text{T}}(AM^2A^{\text{T}})^{-1}\rho^{\text{alias}}$
   This has a block decompotion since each aliased component is the sum of a small number of points in $\mathbf{x} - f$-space

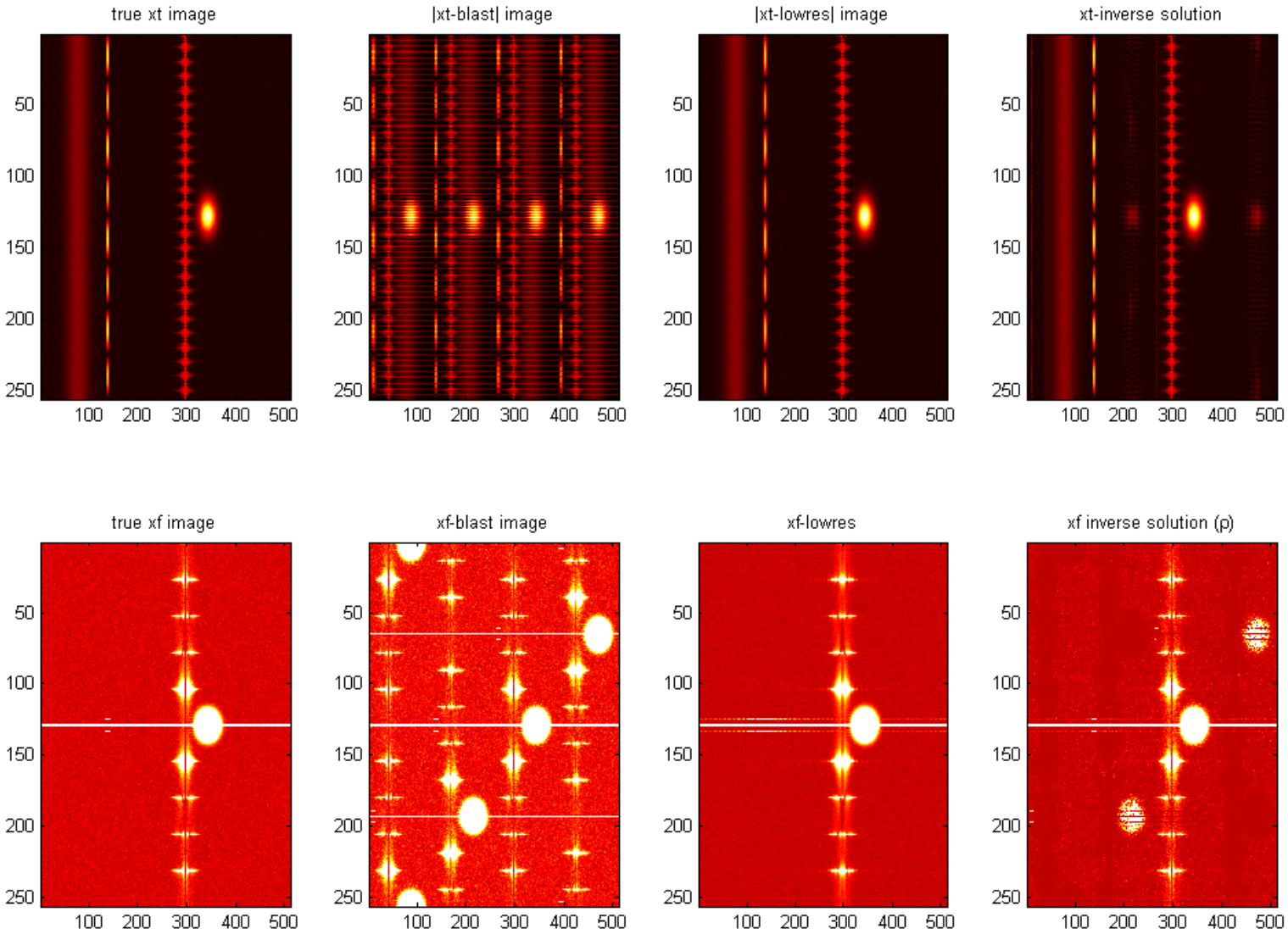
More generally

$$\rho = \rho^{\text{baseline}} + \Gamma_\rho A^{\text{T}}(A\Gamma_\rho A^{\text{T}} + \Gamma_e)^{-1}\rho^{\text{alias}}$$

with Γ_ρ the covariance of a training set, and Γ_e the covariance of noise.

KT-BLAST

1D example

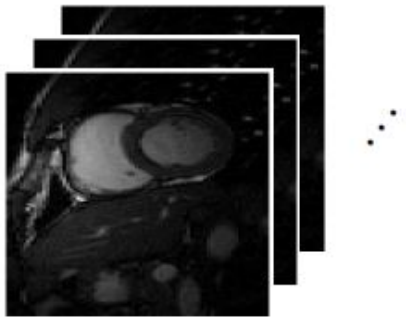


Outline

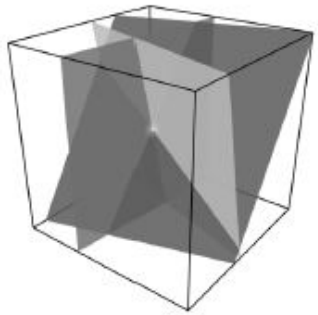
Low-Rank + Sparse Methods

Low Rank Recovery for Dynamic MRI

spatiotemporal signal $\gamma(\mathbf{x}, t)$ exhibits high correlation



Low-dimensional subspace



Low-rank matrix modelling

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma(x_1, t_1) & \dots & \gamma(x_1, t_M) \\ \vdots & \ddots & \vdots \\ \gamma(x_N, t_1) & \dots & \gamma(x_N, t_M) \end{bmatrix}$$

Sparse sampling and reconstruction possible through LR matrix recovery algorithms

$$\mathbf{d} = \mathcal{A}(\mathbf{\Gamma}) + \mathbf{n}$$

Zhao et al, ISBI 2010
Haldar & Liang, ISBI 2011
Gao et al, ISMRM 2012

Low-Rank + Sparse Methods

Outline

Approaches based on low-rank matrix completion for dynamic MR imaging are based on the formulation of the *Casorati matrix*. Each column represents a vectorized complex-valued MR image $x_n \in \mathbb{C}^{N_x N_y}$ (J. P. Haldar and Z.-P. Liang, 2010), so that

$$X = [x_1, \ldots, x_{N_t}] \in \mathbb{C}^{N_x N_y \times N_t}.$$

- In dynamic MR imaging, the Casorati matrix X is very likely to be approximately low-rank, where only a few singular values are significant, because of the high correlation between each images.
- the property of being low-rank for matrices can be seen as the analogous sparsity concept for vectors in CS.
- The finite-dimensional MRI model can be written as

$$y = E(X) + \boldsymbol{n}$$

where now $E : \mathbb{C}^{N_x N_y \times N_t} \rightarrow \mathbb{C}^P$ $(P \ll N_x N_y \times N_t)$ and $X \in \mathbb{C}^{N_x N_y \times N_t}$ represents the matrix to estimate.

Low-Rank + Sparse Methods

Low Rank and Nuclear Norm

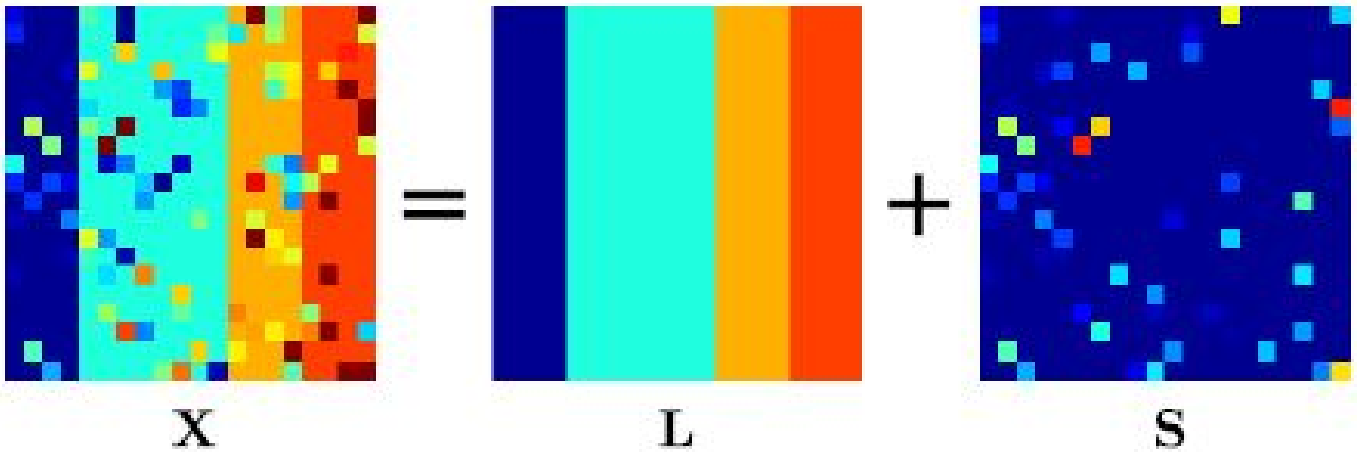
- Similarly to the L_0 norm in CS reconstruction, rank minimization subject to a data fidelity term becomes computationally intractable as the dimension of the problem increases.
- $\Rightarrow$ Change rank penalty to the *nuclear norm*, and relax the equality constraint.
- Nuclear norm (trace norm or Schatten p-norm with p = 1) is the convex envelope of the rank operator (B. Recht, M. Fazel, and P. A. Parrilo, 2010)

$$||\mathbf{X}||_* = \sum_i \sigma_i(\mathbf{X})$$

- In its Lagrangian form, this leads to a nuclear norm regularized linear least squares problem which can be solved efficiently using proximal gradient method (K. C. Toh and S. Yun, 2010),
- it is also possible to solve variants of the rank minimization problem without the use of nuclear norm, for example based on PowerFactorization (J. P. Haldar and D. Hernando, 2009).

Low-Rank + Sparse Methods

Combined Method



- low-rank-plus-sparsity methods (Lingala *et.al.* 2011, Zhao *et.al.* 2012) formulate the minimization problem as,

$$\min_{\mathbf{X}} \left[\frac{1}{2} ||y - E(\mathbf{X})||^2 + \alpha \psi^{\text{rank}}(\mathbf{X}) + \beta \phi^{\text{sparse}}(\mathbf{X}) \right]$$

- $\boldsymbol{k} - t$-SLR : $\psi^{\text{rank}}(\mathbf{X}) \rightarrow ||\mathbf{X}||_p^p$ (non-convex Shatten p-norm), and $\phi^{\text{sparse}} \rightarrow \text{TV}(\mathbf{X})$

Low-Rank + Sparse Methods

Other work

- [Gao PMB'11] robust RPCA for dynamic CT with framelet as sparsifying transform

$$(X_1, X_2) = \underset{(X_1,X_2)}{\arg\min} ||A(X_1 + X_2) - Y||^2 + \lambda_* ||X_1||_* + \lambda_1 ||W X_2||_1$$

- [Lingala TMI'11] uses Schatten-p norm with p=0.1 (non convex) and total variation as sparsifying transform

$$\mathbf{\Gamma}^* = \arg\min_{\mathbf{\Gamma}} \|\mathcal{A}(\mathbf{\Gamma}) - \mathbf{b}\|^2 + \lambda_1 \varphi(\mathbf{\Gamma}) + \lambda_2 \psi(\mathbf{\Gamma})$$

- [Zhao TMI'12] uses partial separability model (to capture spatio temporal correlation) and FT in time as sparsifying transform. Implicitly constraint the rank through the decomposition C=UV with C the Casorati matrix)

$$\hat{\mathbf{U}}_s = \arg \min_{\mathbf{U}_s \in \mathbb{C}^{N \times L}} \|\mathbf{d} - \Omega(\mathbf{F}_s \mathbf{U}_s \mathbf{V}_t)\|_2^2 + \lambda \|\mathrm{vec}(\mathbf{U}_s \mathbf{V}_f)\|_1$$

Outline

Robust Principle Component Analysis

- RPCA decomposes a matrix into low-rank and sparse

$$\min_{\mathbf{L},\mathbf{S}} [|\mathbf{L}|_* + \lambda_\rho |\mathbf{S}|_1] \quad s.t. \, \mathbf{X} = \mathbf{L} + \mathbf{S}$$

- uses ADMM with

$$\begin{aligned} \mathcal{S}_\tau(\mathbf{Y}) &= \text{sgn}(\mathbf{Y})\max(|\mathbf{Y}| - \tau, 0) \\ \mathcal{D}(\mathbf{Y}) &= \mathbf{U}\mathcal{S}_\tau(\mathbf{W})\mathbf{V}^{\mathrm{T}} \end{aligned}$$

Algorithm 1 Robust PCA [16], [31]

Input: $\mathbf{X} \in \mathbb{R}^{N_x N_y \times N_t}$, decomposition parameter $\lambda_\rho > 0$

Initialize: $\mathbf{S}^0 = \mathbf{Z}^0 = 0$, $k = 0$, $\delta = \frac{1}{4}\frac{N_x N_y N_t}{\|\mathbf{X}\|_1}$

while stopping criterion is not met **do**

$$\mathbf{L}^{k+1} \leftarrow \mathcal{D}_{\delta^{-1}}(\mathbf{X} - \mathbf{S}^k + \delta^{-1}\mathbf{Z}^k)$$
$$\mathbf{S}^{k+1} \leftarrow \mathcal{S}_{\lambda_\rho/\delta}(\mathbf{X} - \mathbf{L}^{k+1} + \delta^{-1}\mathbf{Z}^k)$$
$$\mathbf{Z}^{k+1} \leftarrow \mathbf{Z}^k + \delta(\mathbf{X} - \mathbf{L}^{k+1} - \mathbf{S}^{k+1})$$

end while

Output: $\hat{\mathbf{L}}$, $\hat{\mathbf{S}}$

Robust Principle Component Analysis

$\boldsymbol{k}-t$-RPCA, (Trémoulhéac, Dikaios, Atkinson, A, 2014)

- An ADMM algorithm for undersampled $\boldsymbol{k}-t$ data with low rank plus sparsity constraint.
- Fourier transform in time is used for sparsifying
-

$$\min_{\mathbf{P},\mathbf{Q},\mathbf{L},\mathbf{S}} \left[\frac{1}{2}||y - E(\mathbf{L}+\mathbf{S})||^2 + \mu\left(|\mathbf{P}|_* + \lambda_\rho|\mathbf{Q}|_1\right)\right] \quad s.t. \begin{cases} \mathbf{P} & = \mathbf{L} \\ \mathbf{Q} & = \mathcal{F}\mathbf{S} \end{cases}$$

- Associated augmented Lagrangian function

$$\begin{aligned} \mathcal{L}_A &= \frac{1}{2}||y - E(\mathbf{L}+\mathbf{S})||^2 + \mu|\mathbf{P}|_* + \mu\lambda_\rho|\mathbf{Q}|_1 \\ &+ \frac{\delta_1}{2}||\mathbf{L} + \delta_1^{-1}\mathbf{Z}_1 - \mathbf{P}||_2^2 + \frac{\delta_2}{2}||\mathcal{F}\mathbf{S} + \delta_2^{-1}\mathbf{Z}_2 - \mathbf{Q}||_2^2 \end{aligned}$$

minimised over $\mathbf{P}, \mathbf{Q}, \mathbf{L}, \mathbf{S}$ seperately, followed by updating of Lagrangian variables $\mathbf{Z}_1, \mathbf{Z}_2$.

Robust Principle Component Analysis

$k-t$-RPCA algorithm

Algorithm 2 k-t **RPCA**

Input: (k-t)-space samples y, regularization and decomposition parameters $\mu, \lambda_\rho > 0$

Initialize: $\mathbf{X}^0 = \mathbf{L}^0 = \mathbf{E}^H\mathbf{y}$, $\mathbf{S}^0 = \mathbf{Z}_i^0 = 0$, $k = 0$

while stopping criterion is not met **do**

$$\mathbf{P}^{k+1} \leftarrow \mathcal{D}_\mu(\mathbf{X}^k - \mathbf{S}^k + \mathbf{Z}_1^k)$$
$$\mathbf{Q}^{k+1} \leftarrow \mathcal{S}_{\mu\lambda_\rho}(\mathbf{F}_t(\mathbf{X}^k - \mathbf{L}^k) + \mathbf{Z}_2^k)$$
$$\mathbf{L}^{k+1} \leftarrow (\mathbf{E}^H\mathbf{E})^{-1}(\mathbf{E}^H\mathbf{y} + \mathbf{P}^k - \boldsymbol{\Delta}_1^k - \mathbf{E}^H\mathbf{E}\mathbf{S}^k)$$
$$\mathbf{S}^{k+1} \leftarrow (\mathbf{E}^H\mathbf{E})^{-1}(\mathbf{E}^H\mathbf{y} + \mathbf{F}_t^H(\mathbf{Q}^k - \boldsymbol{\Delta}_2^k) - \mathbf{E}^H\mathbf{E}\mathbf{L}^{k+1})$$
$$\mathbf{Z}_1^{k+1} \leftarrow \mathbf{Z}_1^k + \mathbf{L}^{k+1} - \mathbf{P}^{k+1}$$
$$\mathbf{Z}_2^{k+1} \leftarrow \mathbf{Z}_2^k + \mathbf{F}_t(\mathbf{S}^{k+1}) - \mathbf{Q}^{k+1}$$
$$\mathbf{X}^{k+1} \leftarrow \mathbf{L}^{k+1} + \mathbf{S}^{k+1}$$

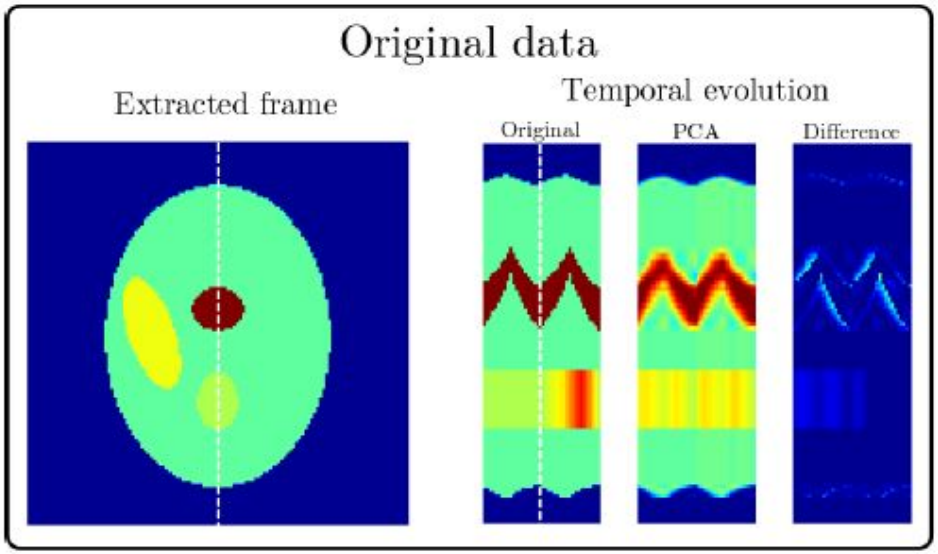
end while

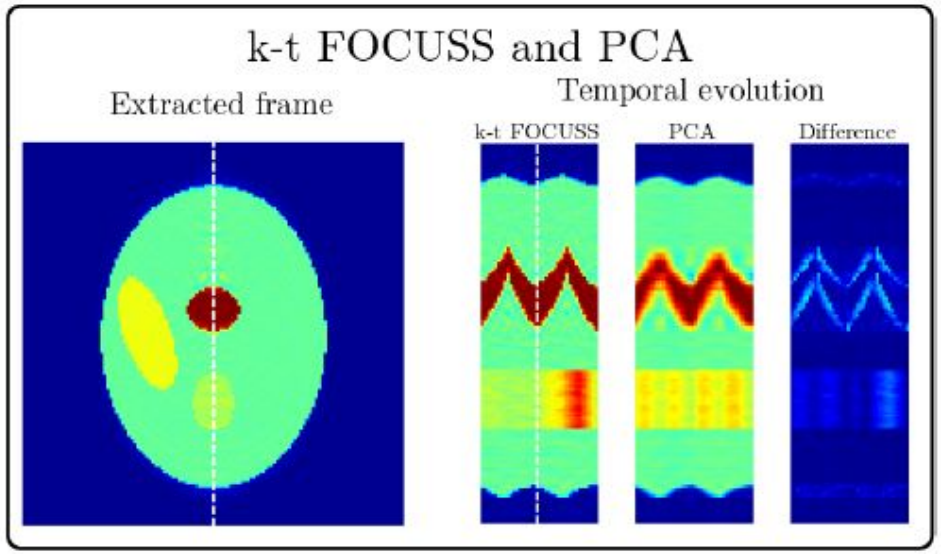
Output: $\hat{\mathbf{L}}$, $\hat{\mathbf{S}}$

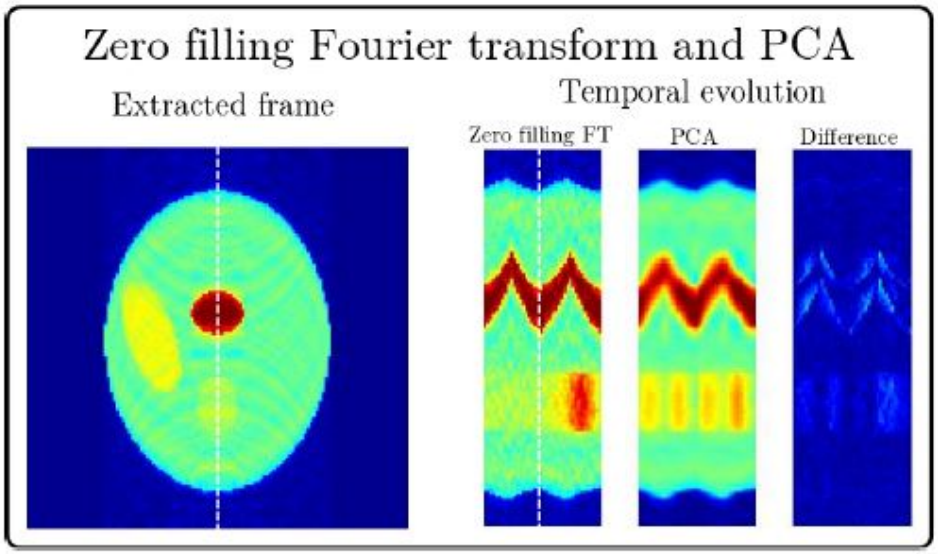
Outline

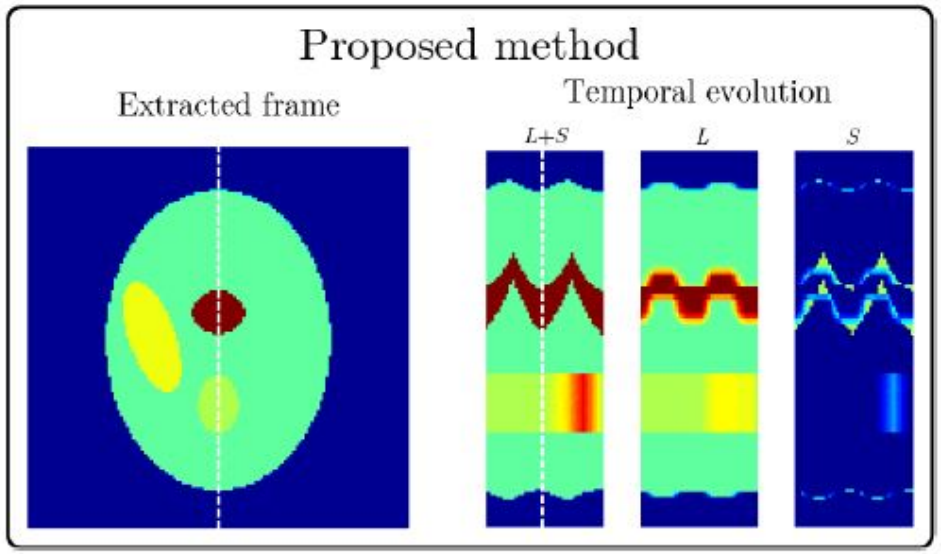
Results

Numerical Phantom





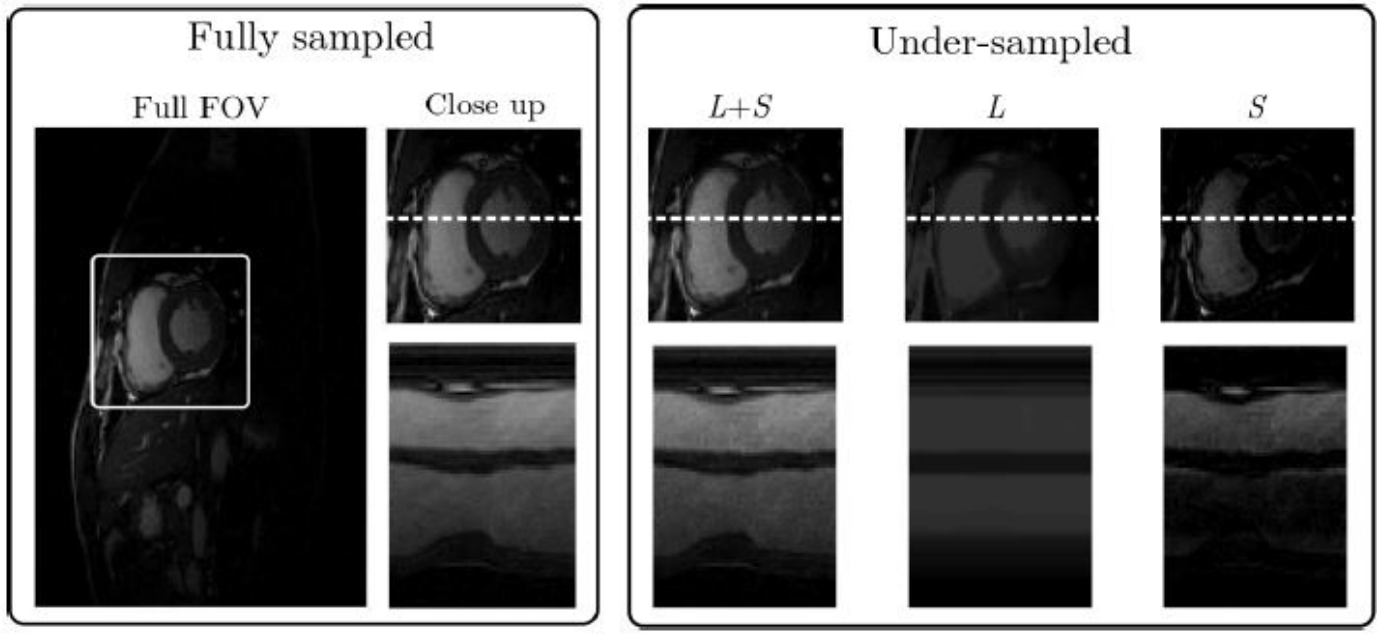




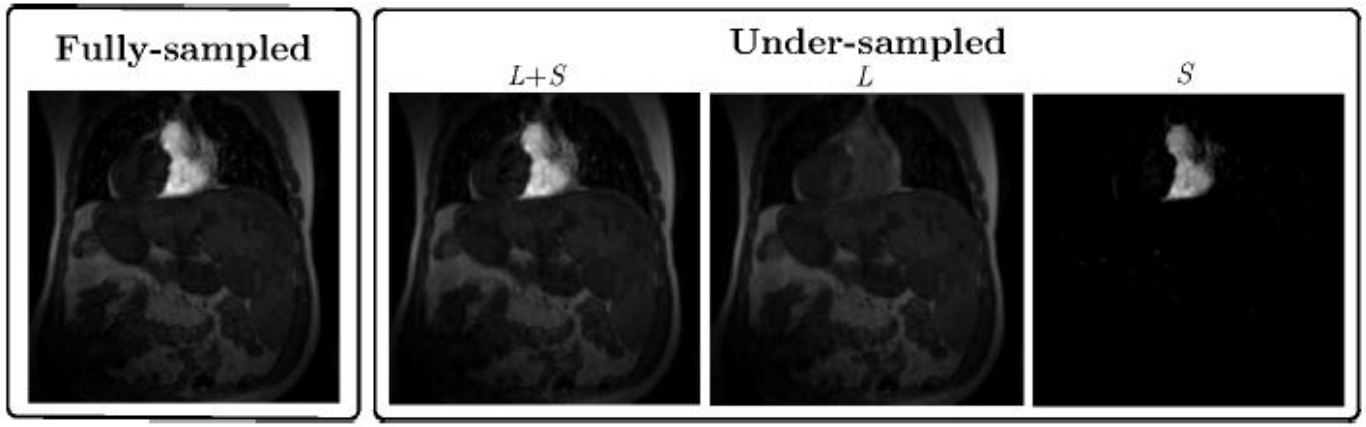
Results

MRI data

Cardiac cine MRI



DCE MRI



Results

Quantitative Results

	Cartesian sampling		Pseudo-radial sampling	
	Reconstruction error in dB (rel. error)	Regularization parameter(s)	Reconstruction error in dB (rel. error)	Regularization parameter(s)
Zero-filled FT	14.3 (3.758291e-02)	N/A	15.6 (2.778515e-02)	N/A
Sliding window	16.7 (2.155588e-02)	N/A	17.2 (1.897659e-02)	N/A
k-t FOCUSS	20.7 (8.558858e-03)	0.1	22.7 (5.419496e-03)	0.01
k-t SLR	24.2 (3.817756e-03)	$\alpha = 0$, $\beta = 1$	28.2 (1.511268e-03)	$\alpha = 0$, $\beta = 1$
k-t RPCA	21.2 (7.639737e-03)	$\mu = 200$, $\rho = 0.75$	23.5 (4.505012e-03)	$\mu = 200$, $\rho = 0.75$

TABLE I

QUANTITATIVE RESULTS FOR NUMERICAL PHANTOM WITH 10-FOLD ACCELERATION.

	Cartesian sampling		Pseudo-radial sampling	
	Reconstruction error in dB (rel. error)	Regularization parameter(s)	Reconstruction error in dB (rel. error)	Regularization parameter(s)
Zero-filled FT	8.79 (1.321385e-01)	N/A	***	N/A
Sliding window	12.5 (5.659558e-02)	N/A	***	N/A
k-t FOCUSS	15.7 (2.676942e-02)	0.01	18.1 (1.554241e-02)	0
k-t SLR	15.7 (2.662888e-02)	$\alpha = 100$, $\beta = 0.01$	17.8 (1.653227e-02)	$\alpha = 100$, $\beta = 0.01$
k-t RPCA	15.6 (2.768693e-02)	$\mu = 200$, $\rho = 1.75$	19.1 (1.233944e-02)	$\alpha = 100$, $\beta = 1.75$

TABLE II

QUANTITATIVE RESULTS FOR FREE BREATHING REAL CARDIAC MRI DATA WITH ABOUT 10-FOLD ACCELERATION (CARTESIAN SAMPL

Results

comparison of methods

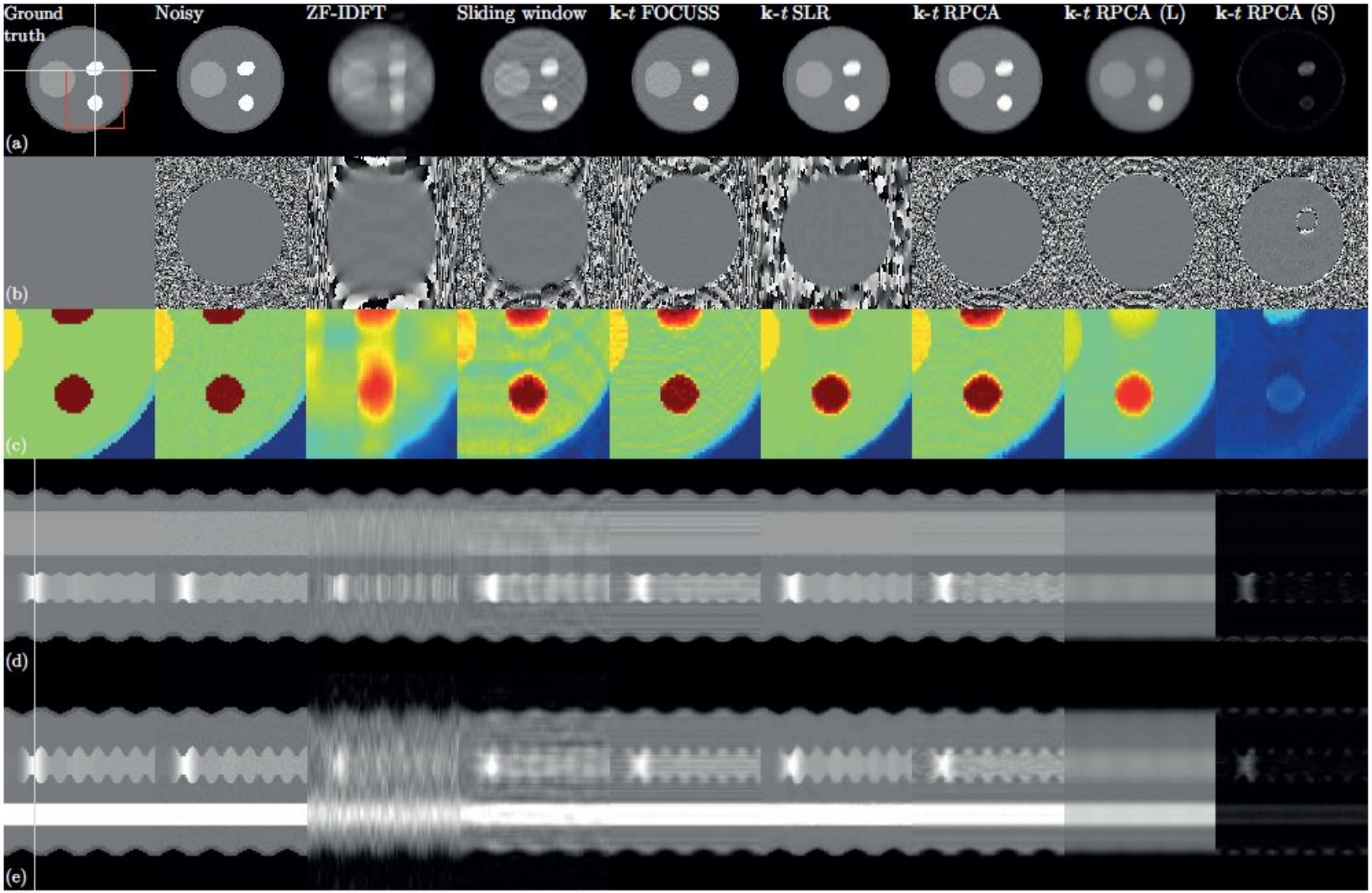


Fig. 5. Qualitative results for Shepp-Logan based phantom reconstruction for 10-fold acceleration and Cartesian sampling. (a) Magnitude images (b) Phase images (c) Zoom-in magnitude images (corresponding to the red square on the ground truth image) (d) x-t temporal profiles and (e) y-t temporal profiles (according to the dotted lines on the ground truth image). The time frames shown in the first three rows correspond to the frames selected on the dotted lines on the temporal profiles.

Outline

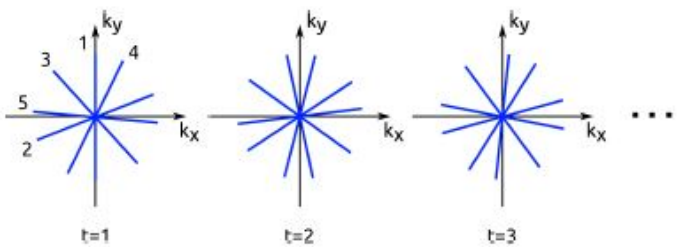
k-t NNAPG: Fast dynamic MRI via nuclear norm minimization and accelerated proximal gradient [ISBI'13]

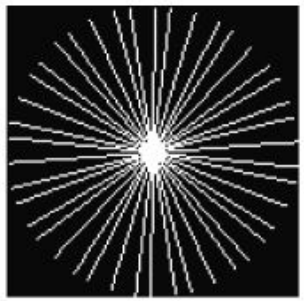
The idea was to find a fast algorithm using only rank prior information and test its efficiency *in the context of under-sampled dynamic MRI*; which so far hasn't been really investigated apart from [Haldar&Liang ISBI'10,Zhao ISBI'10]

The problem was formulated as a convex one implying only nuclear norm regularization

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X} \in \mathbb{C}^{N \times M}} \frac{1}{2} \|\mathcal{A}(\mathbf{X}) - \mathbf{y}\|_2^2 + \mu \|\mathbf{X}\|_*$$

This was solved through an accelerated proximal gradient (gradient step + proximal step for nuclear norm) algorithm using a randomly-rotated golden angle based radial sampling.

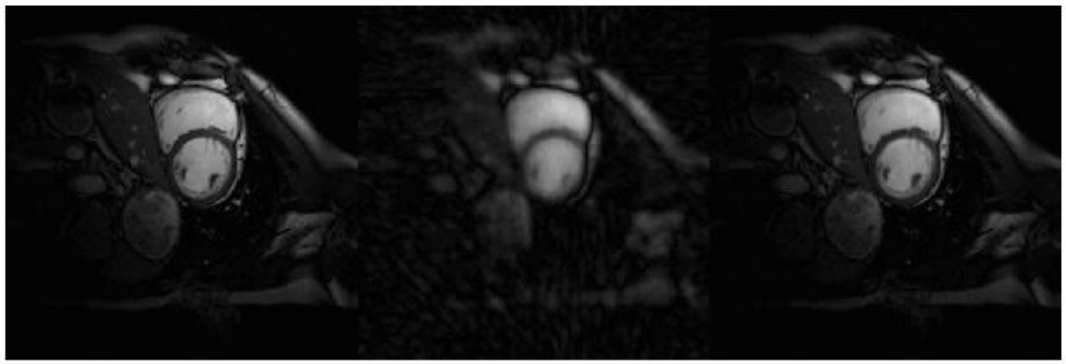




Accelerated Proximal Gradient Method

NNAPG results

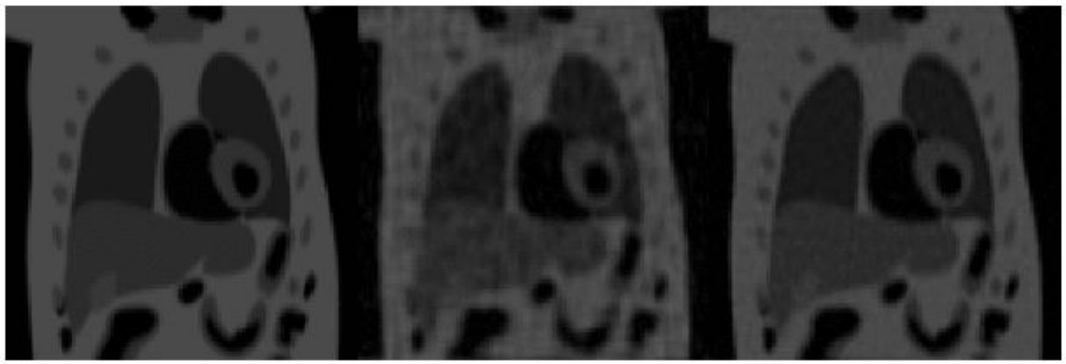
k-t NNAPG: Fast dynamic MRI via nuclear norm minimization and accelerated proximal gradient [ISBI'13]



Cine cardiac MR imaging dataset (breath-hold).

Ground truth / direct (zero-filled) inverse Fourier transform (9.9dB) / k-t NNAPG reconstruction (17.8dB)

8.7-fold acceleration, reconstruction time of **58sec** (2.2GHz/12GB laptop/Matlab non optimised code)



PINCAT phantom with contrast (free-breathing)

Ground truth / direct (zero-filled) inverse Fourier transform (19.4dB) / k-t NNAPG reconstruction (30.2dB)

4.5-fold acceleration, reconstruction time of **22sec** (2.2GHz/12GB laptop/Matlab non optimised code)

Outline

Summary and Outlook

- two methods for reconstruction of low-rank and sparse components from under-sampled dynamic MRI
  1. k-t RPCA, a motion and contrast enhancement separation model for under-sampled dynamic MRI based on low-rank plus sparse decomposition
  2. k-t NNAPG, a low-rank matrix recovery method for accelerated dynamic MRI that is quite efficient and fast in terms of computational time
- Enables faster dynamic MR imaging while providing a flexible separation of motion and contrast enhancement
- Is there an optimal random under-sampling scheme for low-rank methods?
- Extend methods to parallel imaging
- Develop a sparsifying transform for dynamic MR data
- Incorporate motion correction into the reconstruction
- Going to higher dimension: tensor approaches, 3D imaging

Outline

Acknowledgements

- **Collaborators :**
  - *UCL*: David Atkinson, Nikoaos Dikaios, Benjamin Trémoulhéac,
- **Funding**
  - This work was supported by EPSRC grant EP/H046410/1 “Intelligent Imaging: Motion, Form and Function Across Scale”

PhD Advert

Image Reconstruction from In-perfect Data in Photoacoustic Tomography

deadline 1st May

http://prism.ucl.ac.uk/pgadmissions/apply/new?program=RRDCOMSING01&project=19&advert=55

contact me at
m.betcke@ucl.ac.uk