Non-convex optimization for robust multi-view imaging

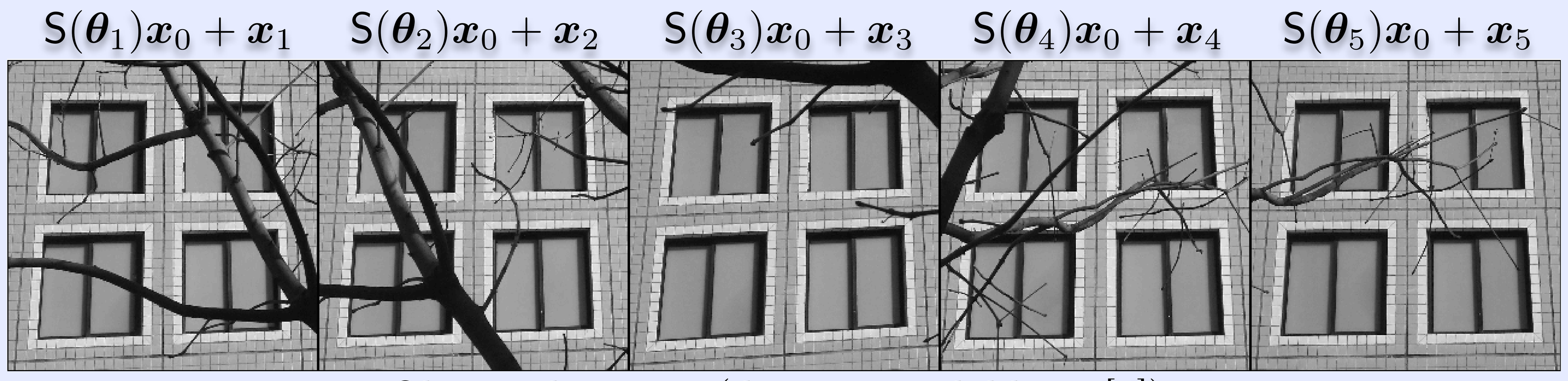
G. Puy and P. Vandergheynst

Institute of Electrical Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

Problem formulation

- A scene $\boldsymbol{x}_0 \in \mathbb{R}^n$ is observed from l viewpoints, providing l noisy measurement vectors $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_l \in \mathbb{R}^m$, $m \leqslant n$.
- The scene is partly occluded by some objects $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_l \in \mathbb{R}^n$.
- The scene undergoes geometric transformations that depend on the position of the observer. We assume that they can be represented by p parameters: $\theta_j \in \mathbb{R}^p, \forall j \in \{1, \ldots, l\}$.

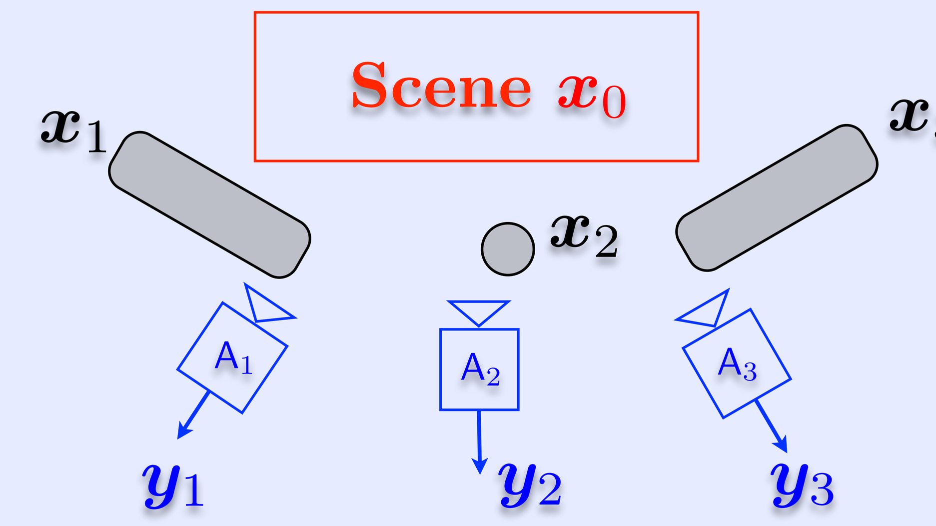
$\mathsf{S}(\boldsymbol{\theta}_1)\boldsymbol{x}_0 + \boldsymbol{x}_1$ $\mathsf{S}(\boldsymbol{\theta}_2)\boldsymbol{x}_0 + \boldsymbol{x}_2$ $\mathsf{S}(\boldsymbol{\theta}_3)\boldsymbol{x}_0 + \boldsymbol{x}_3$ $\mathsf{S}(\boldsymbol{\theta}_4)\boldsymbol{x}_0 + \boldsymbol{x}_4$ $\mathsf{S}(\boldsymbol{\theta}_5)\boldsymbol{x}_0 + \boldsymbol{x}_5$



Observed images (dataset available in [1]).

[1] Peng *et al.*, CVPR, pp. 763-770, 2010.

- The transformed images are estimated using interpolation matrices $\mathsf{S}(\theta_j) \in \mathbb{R}^{n \times n}$, $j = 1, \ldots, l$, built using, e.g., bicubic splines.
- Observation processes are modeled by $\mathsf{A}_1, \ldots, \mathsf{A}_l \in \mathbb{R}^{m \times n}$.
- Our goal is to reconstruct $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_l$, and estimate $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_l$, using $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_l$ as sole information.



Optimization method

- The measurement model satisfies

$$\underbrace{\begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_l \end{bmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{bmatrix} \mathsf{A}_1\mathsf{S}(\theta_1) & \mathsf{A}_1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{A}_l\mathsf{S}(\theta_l) & 0 & \ldots & \mathsf{A}_l \end{bmatrix}}_{\mathsf{A}(\theta)} \underbrace{\begin{bmatrix} \boldsymbol{x}_0 \\ \vdots \\ \boldsymbol{x}_l \end{bmatrix}}_{\boldsymbol{x}} + \begin{bmatrix} \boldsymbol{n}_1 \\ \vdots \\ \boldsymbol{n}_l \end{bmatrix}.$$

- Find a solution to this ill-posed inverse problem by minimizing

$$L(\boldsymbol{x}, \boldsymbol{\theta}) = f(\boldsymbol{x}) + \kappa \left\| \mathsf{A}(\boldsymbol{\theta})\, \boldsymbol{x} - \boldsymbol{y} \right\|_2^2 + i_\Theta(\boldsymbol{\theta}).$$

Prior on $\boldsymbol{x}$, e.g., ℓ_1, TV. Parameters belong to a convex set Θ.

- We use an alternate minimization method inspired by [2], [3]. Generate a sequence of estimates $(\boldsymbol{x}^k, \boldsymbol{\theta}^k)_{k \in \mathbb{N}}$ converging to a critical point $(\boldsymbol{x}^*, \boldsymbol{\theta}^*)$ of L [4].

- Start from $\boldsymbol{x}^0 = \mathbf{0} \in \mathbb{R}^{(l+1)n}$ and $\boldsymbol{\theta}^0 \in \Theta$.
- Step 1: update of the images.

$$\boldsymbol{x}^{k+1} \leftarrow \underset{\boldsymbol{x}}{\operatorname{argmin}}\; L(\boldsymbol{x}, \boldsymbol{\theta}^k) + \frac{\lambda_{\boldsymbol{x}}^k}{2}\, \underbrace{h_\mu\left(\Psi^\intercal(\boldsymbol{x} - \boldsymbol{x}^k)\right)}_{\approx \|\Psi^\intercal(\boldsymbol{x} - \boldsymbol{x}^k)\|_1}.$$

Wavelet tight-frame. Coarse-to-fine scales reconstruction. ($k \leftarrow k+1$)

- Step 2: update of the parameters.

$$\|\mathsf{A}(\boldsymbol{\theta})\boldsymbol{x}^{k+1} - \boldsymbol{y}\|_2^2 = \sum_{1 \leqslant j \leqslant l} \|\mathsf{A}_j \mathsf{S}(\boldsymbol{\theta}_j)\boldsymbol{x}_0^{k+1} + \mathsf{A}_j \boldsymbol{x}_j^{k+1} - \boldsymbol{y}_j\|_2^2 = \sum_{1 \leqslant j \leqslant l} Q_j(\boldsymbol{\theta}_j).$$

For $j = 1, \ldots, l$, minimize a quadratic approx. of Q_j at $\boldsymbol{\theta}_j^k$:

$$\boldsymbol{\theta}_j^{k+1} \leftarrow \underset{\boldsymbol{\theta}_j \in \Theta_j}{\operatorname{argmin}}\; P_j(\boldsymbol{\theta}_j) = \nabla Q_j(\boldsymbol{\theta}_j^k)^\intercal(\boldsymbol{\theta}_j - \boldsymbol{\theta}_j^k) + (\boldsymbol{\theta}_j - \boldsymbol{\theta}_j^k)^\intercal \left[\frac{\mathsf{H}_j^k + 2^i \lambda_{\boldsymbol{\theta}} \mathsf{I}}{2}\right] (\boldsymbol{\theta}_j - \boldsymbol{\theta}_j^k),$$

where i is the smallest integer such that:

$$Q_j(\boldsymbol{\theta}_j^{k+1}) + \lambda_{\boldsymbol{\theta}}/2\; \|\boldsymbol{\theta}_j^{k+1} - \boldsymbol{\theta}_j^k\|_2^2 \leqslant Q_j(\boldsymbol{\theta}_j^k) + P_j(\boldsymbol{\theta}_j^{k+1}).$$

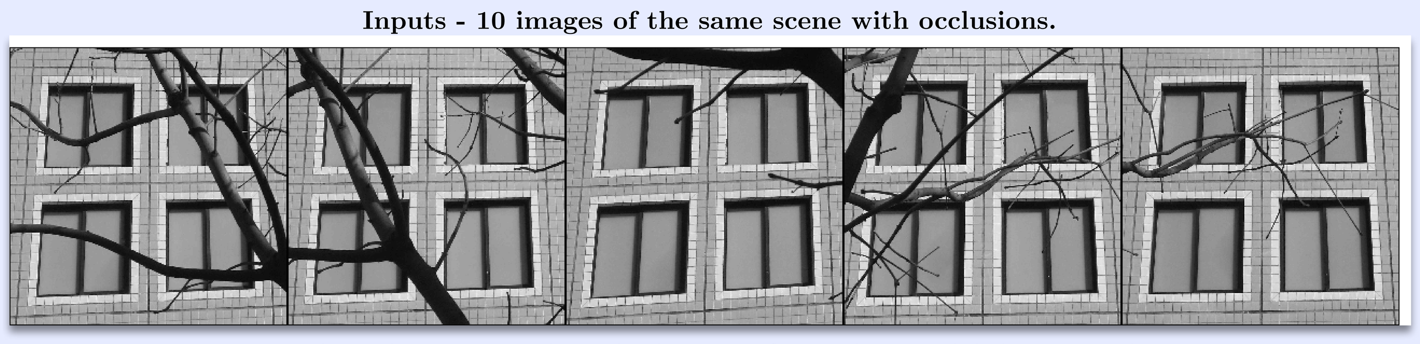
"Sufficient" decrease condition of the objective function

[2] Attouch *et al.*, Mathematics of Operations Research, vol. 35(2), pp. 438-457, 2010. [3] Attouch *et al.*, J. Mathematical programming, 2011. [4] Puy *et al.*, SIAM J. on Imaging Sciences, submitted, 2012.

Applications

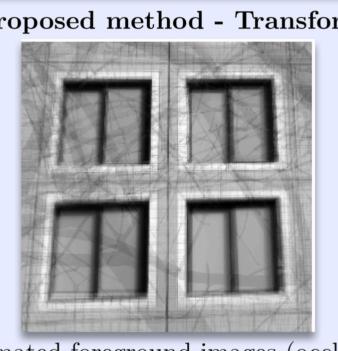
Robust image alignment

Inputs - 10 images of the same scene with occlusions.

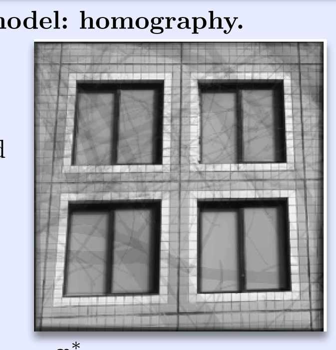


Results using proposed method - Transformation model: homography.

Estimated background image $\boldsymbol{x}_0^*$.

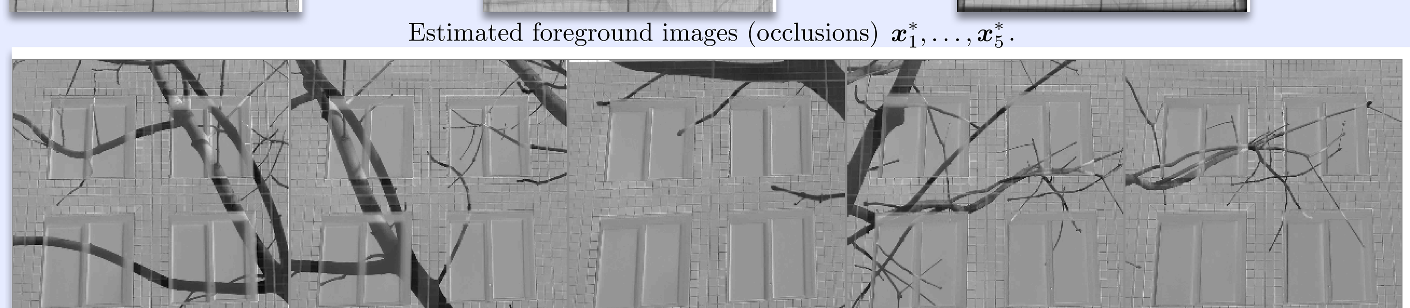


Superposed input images.



Superposed **registered** input images using $\boldsymbol{\theta}^*$.

Estimated foreground images (occlusions) $\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_5^*$.



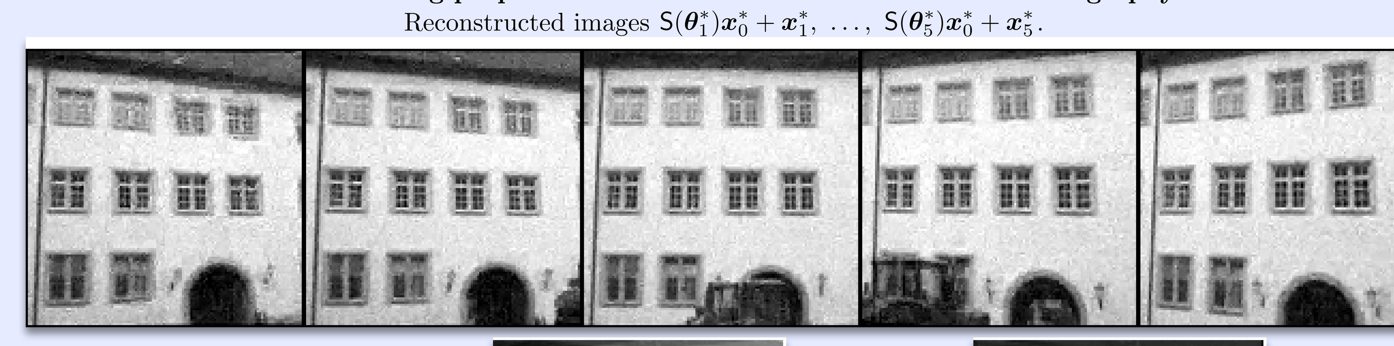
Aligned background images $\mathsf{S}(\boldsymbol{\theta}_1^*)\boldsymbol{x}_0^*, \ldots, \mathsf{S}(\boldsymbol{\theta}_5^*)\boldsymbol{x}_0^*$.

Compressed sensing

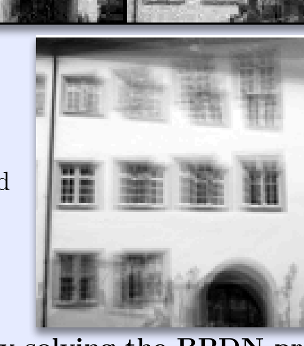
Inputs - $m = 0.3n$ compressed measurements of the 5 images below.

Results using proposed method - Transformation model: homography.

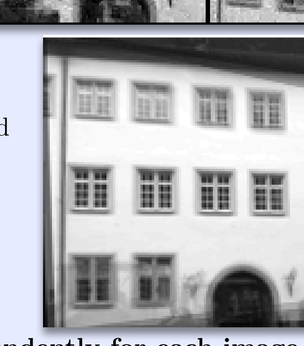
Reconstructed images $\mathsf{S}(\boldsymbol{\theta}_1^*)\boldsymbol{x}_0^* + \boldsymbol{x}_1^*, \ldots, \mathsf{S}(\boldsymbol{\theta}_5^*)\boldsymbol{x}_0^* + \boldsymbol{x}_5^*$.



Estimated background image $\boldsymbol{x}_0^*$.



Superposed ground-truth images.



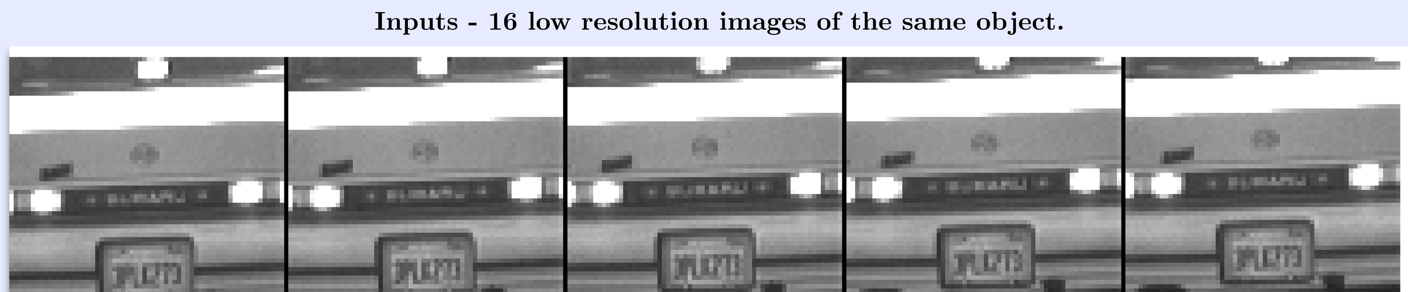
Superposed **registered** ground-truth images using $\boldsymbol{\theta}^*$.

Results obtained by solving the BPDN problem independently for each image.

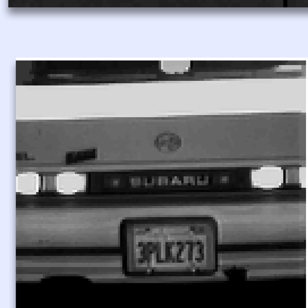
Castle-R20 dataset available at cvlab.epfl.ch

Super-resolution

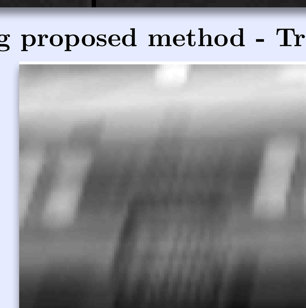
Inputs - 16 low resolution images of the same object.



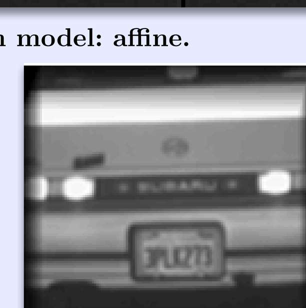
Results using proposed method - Transformation model: affine.



Estimated high resolution background image $\boldsymbol{x}_0^*$.



Superposed input images.



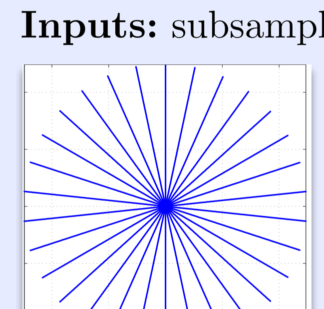
Superposed **registered** input images using $\boldsymbol{\theta}^*$.

Dataset available at users.soe.ucsc.edu/~milanfar/software/sr-datasets.html

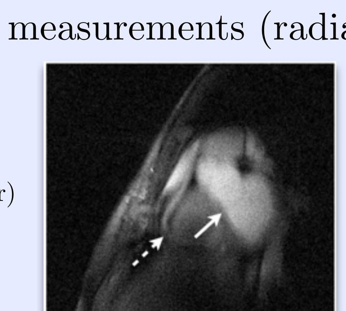
Free breathing coronary MRI

Goal: obtain a high-resolution image of the heart.

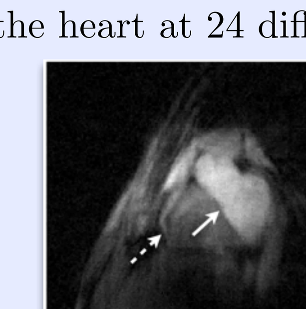
Inputs: subsampled Fourier measurements (radial sampling) of the heart at 24 different positions.



k-space (Fourier) sampling



Reconstructed image using a usual technique without correction of heart motion.



Reconstructed image with **proposed method**, which **automatically corrects** heart motion.

Joint work with G. Bonanno and M. Stuber.

Conclusion

- The method jointly reconstructs misaligned images, while estimating the transformation between them.
- Extension to more challenging scene structures thanks to the use of parametric elastic transformations.