Light Field Compressive Sensing in Camera Arrays

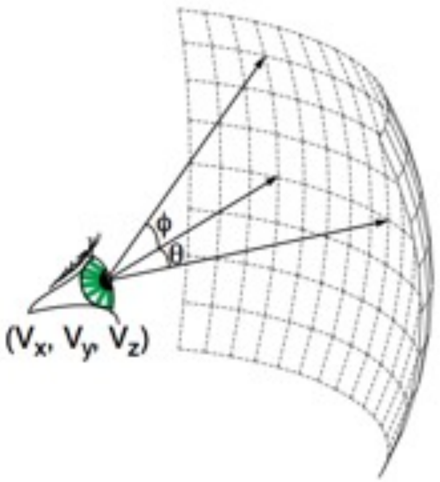
Mahdad Hosseini Kamal

Joint work with:

M. Golbabaee & P. Vandergheynst

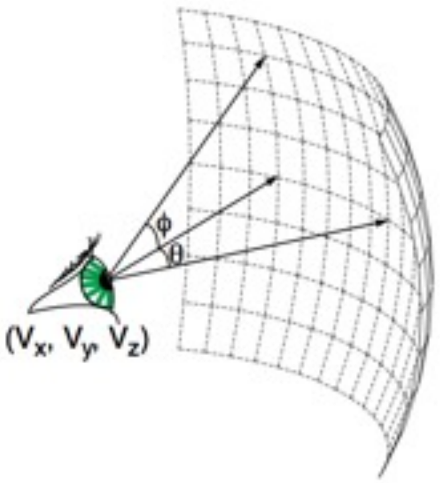
Light Field

- The amount of light traveling in every direction through every point in space.
- Light field can be captured using many cameras having different view points.
- Light field rendering does not need any geometrical information.



Light Field

- The amount of light traveling in every direction through every point in space.
- Light field can be captured using many cameras having different view points.
- Light field rendering does not need any geometrical information.
- Applications:
    - Generating a new view [1]

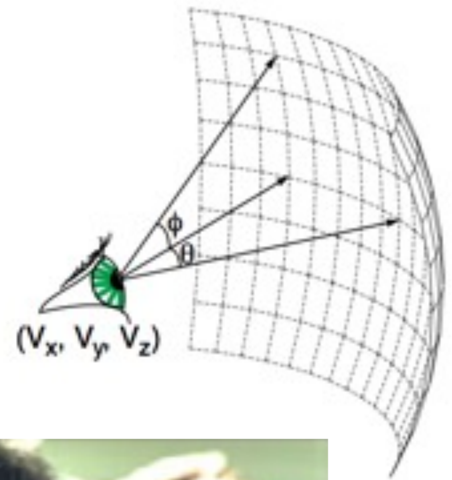


[1] S.J. Gortler et al., “The Lumigraph”, ACM SIGGRAPH 96,

M. Levoy et al., “Light Field Rendering”, ACM SIGGRAPH 96

Light Field

- The amount of light traveling in every direction through every point in space.
- Light field can be captured using many cameras having different view points.
- Light field rendering does not need any geometrical information.
- Applications:
  - Generating a new view
  - Digital refocusing[2]



[2] R. Ng et al., “Light Field Photography with a Hand-held Plenoptic Camera”,CSTR 05

Light Field Cameras

- Light field is captured by an array of cameras with overlapping fields of view.
- Large amount of data.
- Highly correlated images.
- It is necessary to employ a compression scheme.

Compressive Sensing

- A sparse signal can be reconstructed from a relatively small number of linear measurements.

$$x = \Phi\alpha \qquad y = \mathcal{A}\Phi\alpha \qquad m \ll n$$

- $\mathcal{A}$: measurement matrix, for example Random Convolution.[1]
- x : K-sparse in an orthobasis Φ → $m = O(K \log n/K)$
- Recovery takes explicit advantage of sparsity.

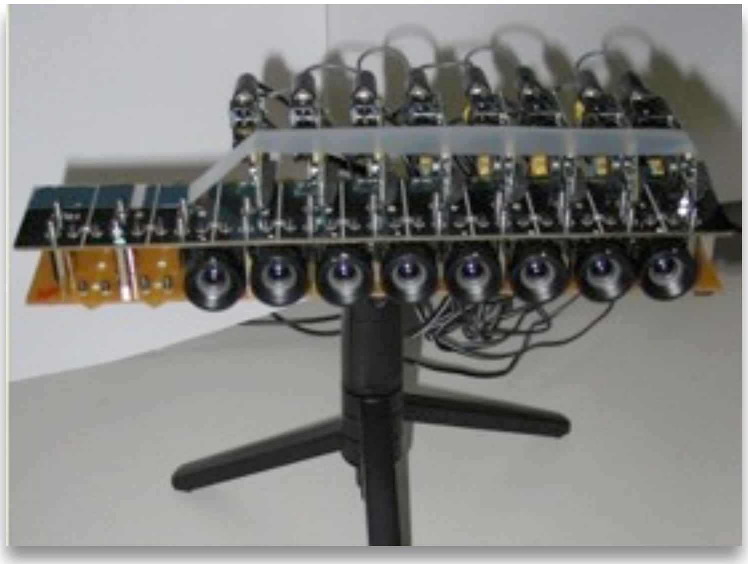
$$\text{argmin } ||\alpha||_1 \quad \textit{subject to} \quad ||y - \mathcal{A}\Phi\alpha||_2 \leq \epsilon$$

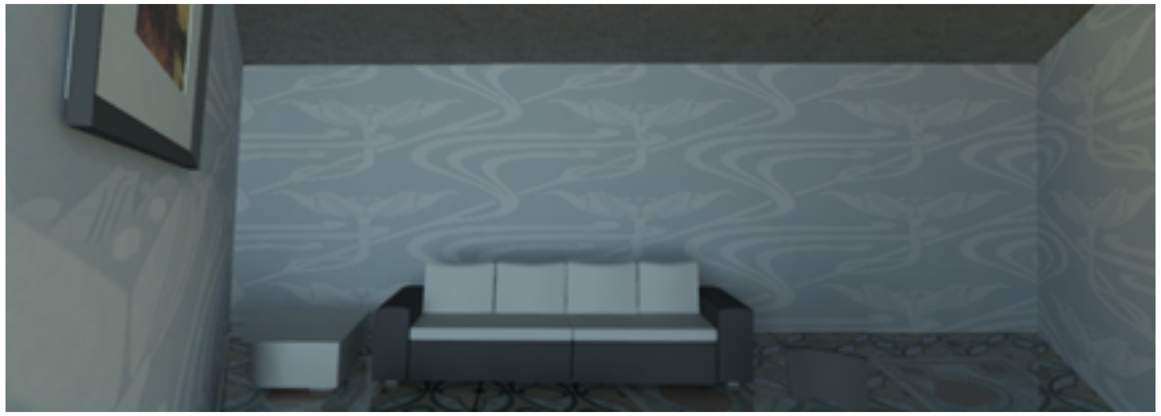
[1] J. Romberg, "Compressive Sensing by Random Convolution", SIAM SIIMS, 09,

L. Jacques, et al., "CMOS Compressed Imaging by Random Convolution", ICASSP 09

Problem Definition

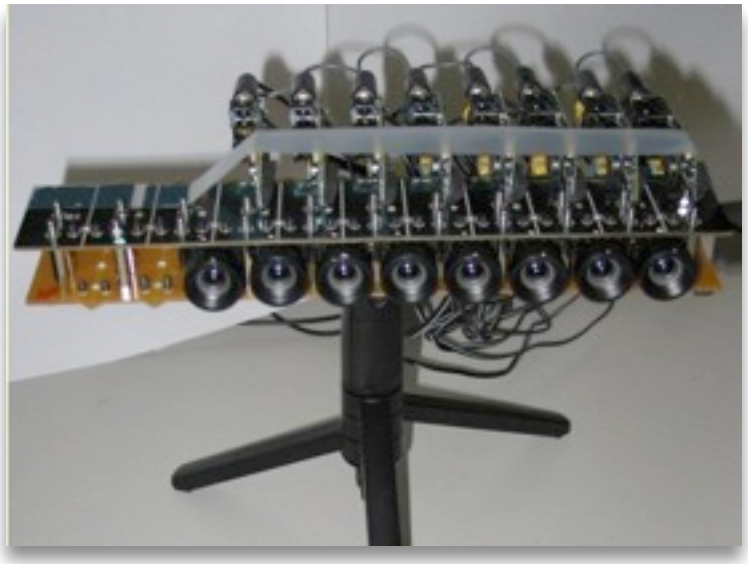
- A 1D array of 40 cameras to capture the light fields.
- Each camera observes a part of the scene.
- Captured images are highly correlated.
- Using a dictionary to exploit the redundancy.

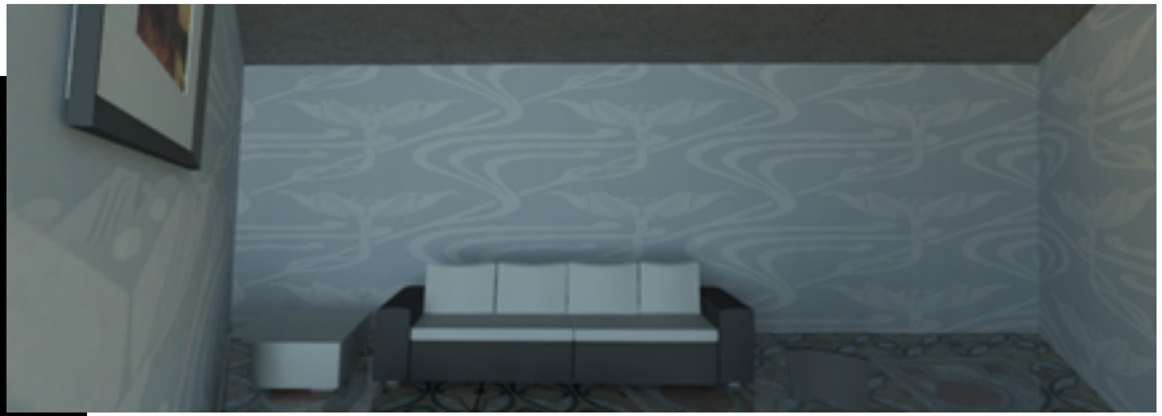




Problem Definition

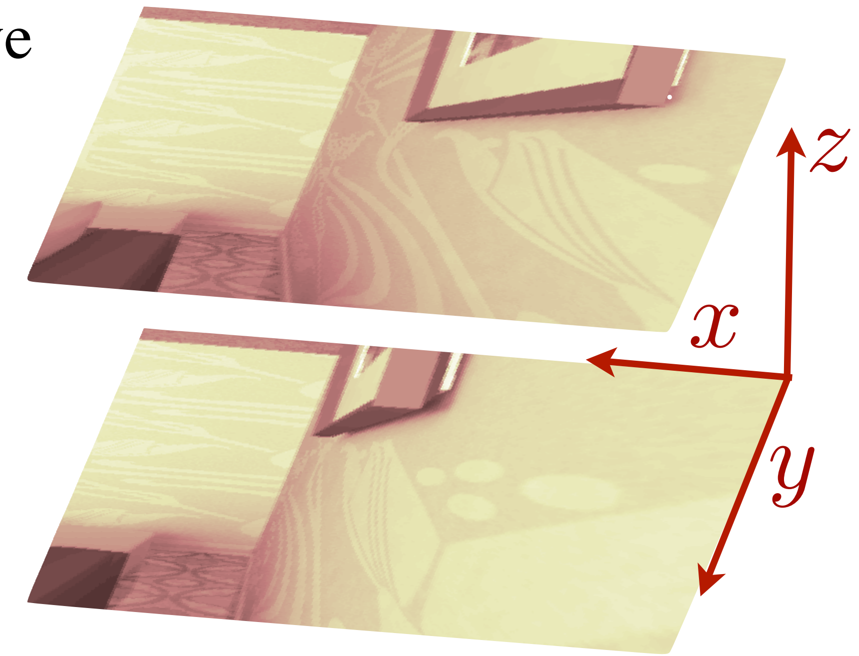
- A 1D array of 40 cameras to capture the light fields.
- Each camera observes a part of the scene.
- Captured images are highly correlated.
- Using a dictionary to exploit the redundancy.





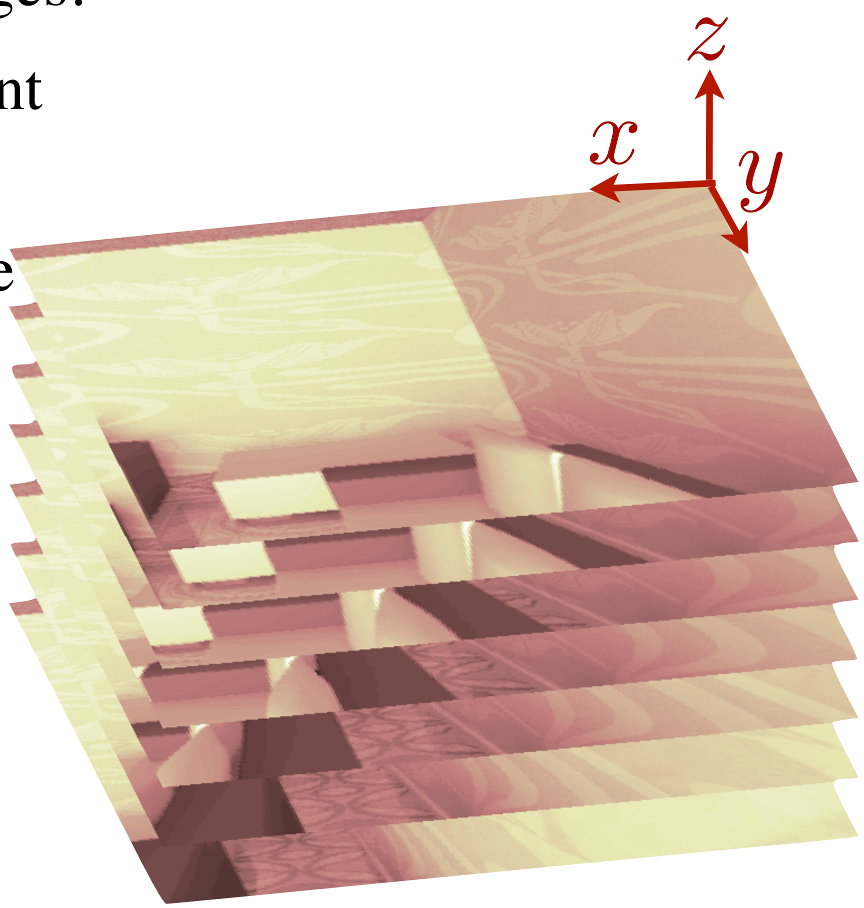
Epipolar Plane Image

- Stack the images observed by each camera on top of each other to make an image volume $\mathcal{X}$.
- Projection of scene objects has linear movement in images.
- Objects at different depth levels have different movement speeds.
- Slices of the image volume along moving direction have stair-like shapes, called Epipolar plane images (EPI).



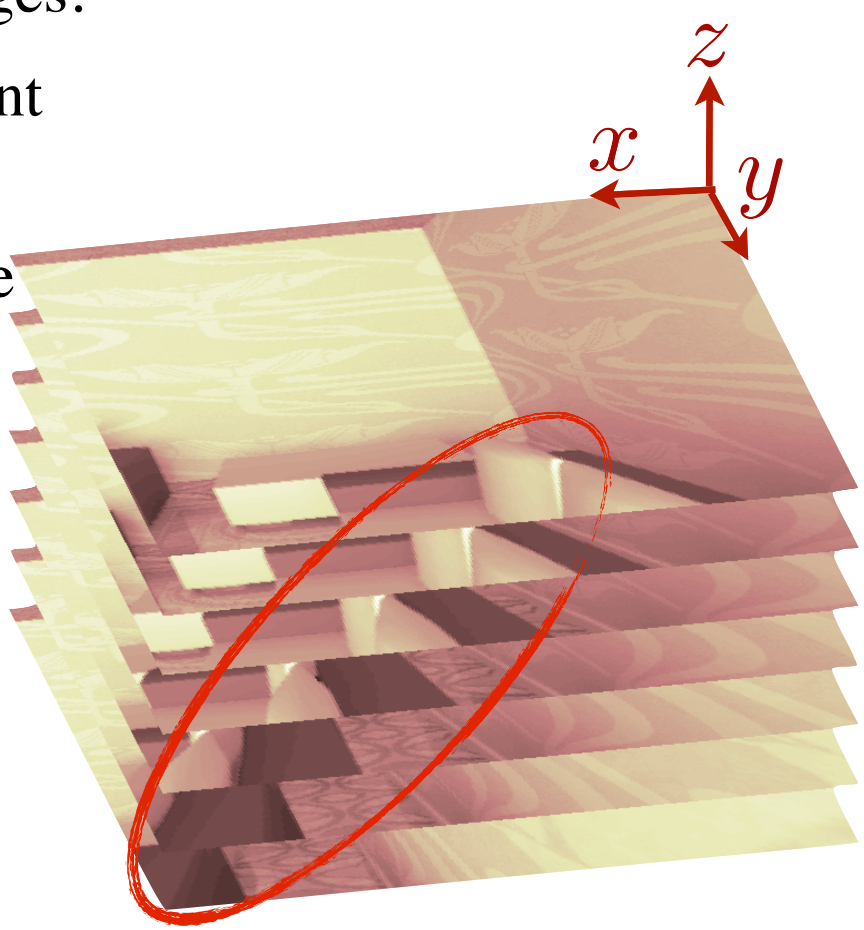
Epipolar Plane Image

- Stack the images observed by each camera on top of each other to make an image volume $\mathcal{X}$.
- Projection of scene objects has linear movement in images.
- Objects at different depth levels have different movement speeds.
- Slices of the image volume along moving direction have stair-like shapes, called Epipolar plane images (EPI).



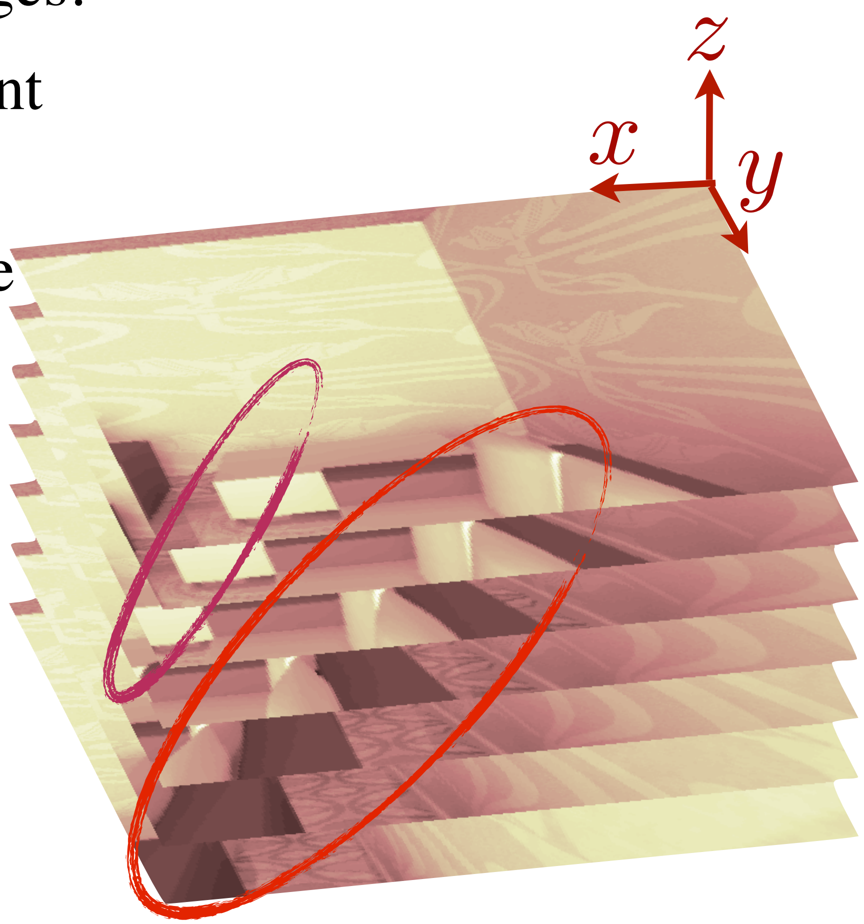
Epipolar Plane Image

- Stack the images observed by each camera on top of each other to make an image volume $\mathcal{X}$.
- Projection of scene objects has linear movement in images.
- Objects at different depth levels have different movement speeds.
- Slices of the image volume along moving direction have stair-like shapes, called Epipolar plane images (EPI).



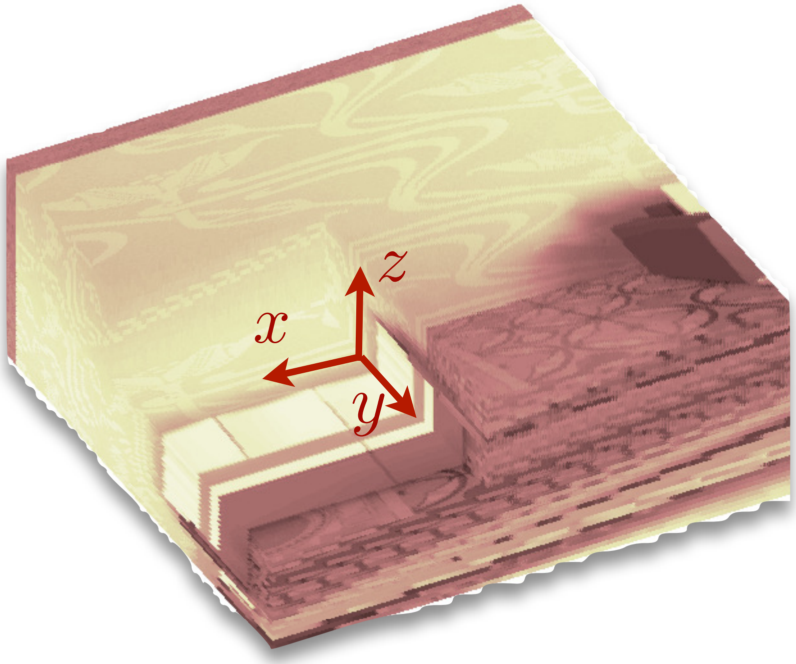
Epipolar Plane Image

- Stack the images observed by each camera on top of each other to make an image volume $\mathcal{X}$.
- Projection of scene objects has linear movement in images.
- Objects at different depth levels have different movement speeds.
- Slices of the image volume along moving direction have stair-like shapes, called Epipolar plane images (EPI).



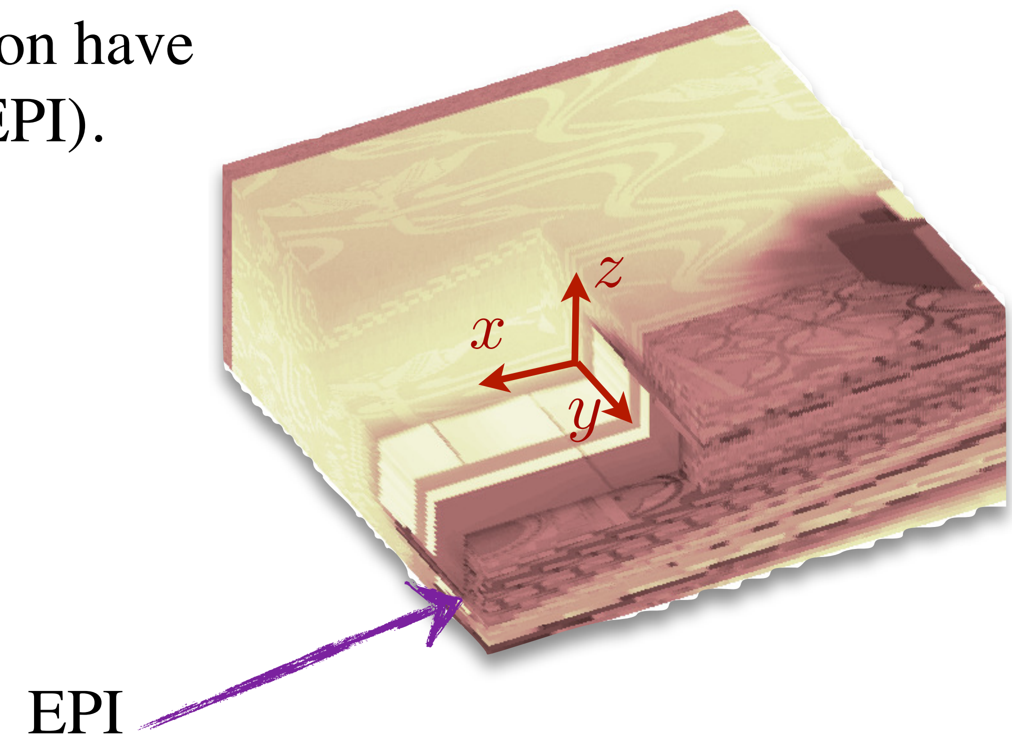
Epipolar Plane Image

- Stack the images observed by each camera on top of each other to make an image volume $\mathcal{X}$.
- Projection of scene objects has linear movement in images.
- Objects at different depth levels have different movement speeds.
- Slices of the image volume along moving direction have stair-like shapes, called Epipolar plane images (EPI).



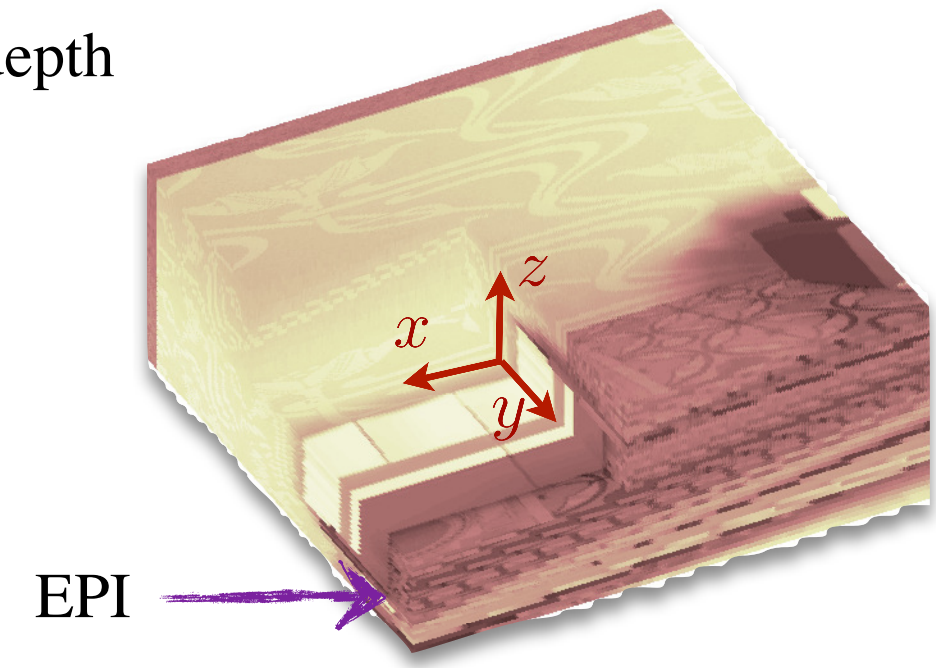
Epipolar Plane Image

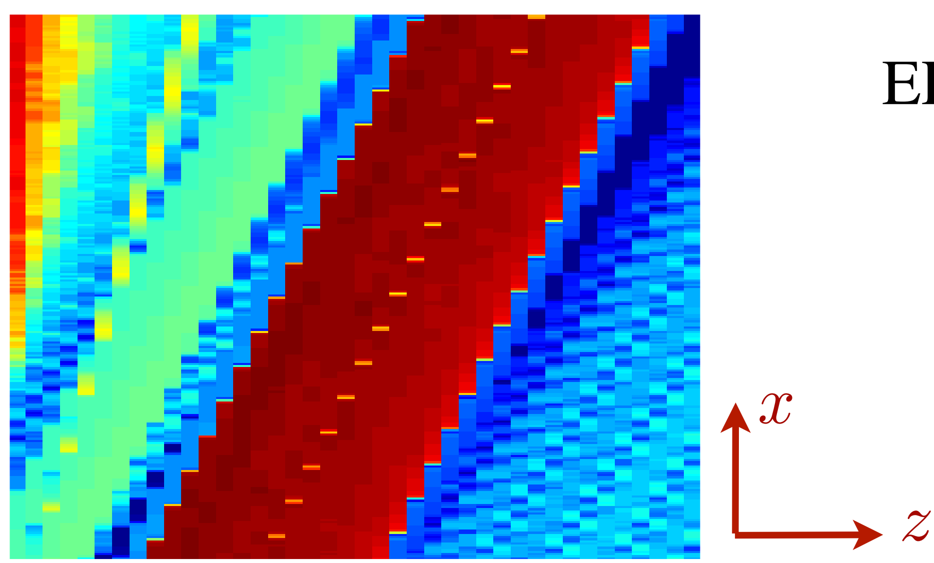
- Stack the images observed by each camera on top of each other to make an image volume $\mathcal{X}$.
- Projection of scene objects has linear movement in images.
- Objects at different depth levels have different movement speeds.
- Slices of the image volume along moving direction have stair-like shapes, called Epipolar plane images (EPI).



Epipolar Plane Image

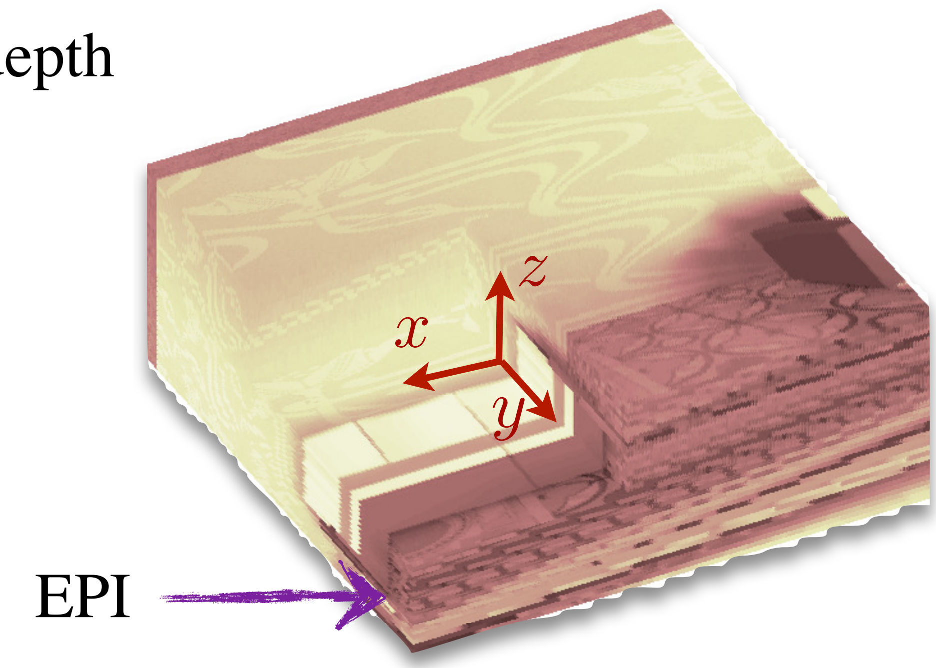
- Each line represents the path a pixel travels in different cameras.
- Different line slopes are consequence of different depth levels.

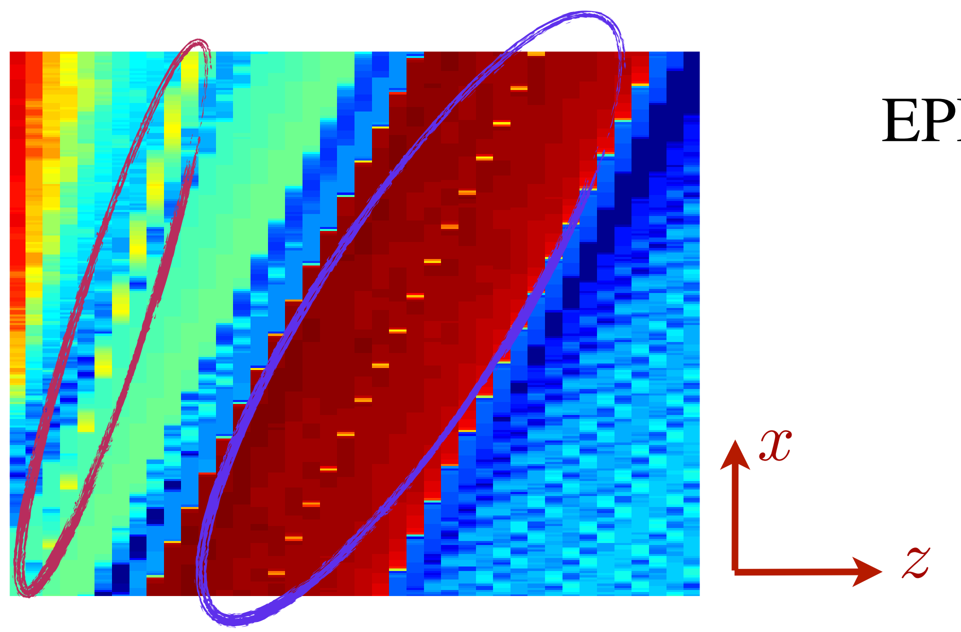




Epipolar Plane Image

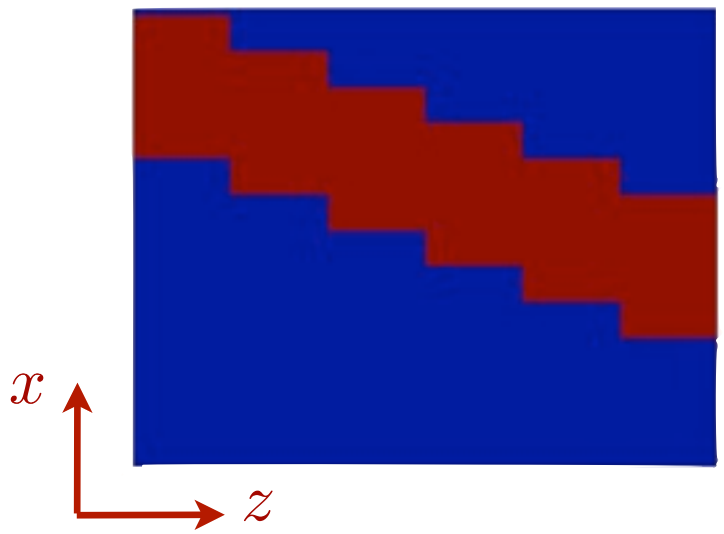
- Each line represents the path a pixel travels in different cameras.
- Different line slopes are consequence of different depth levels.

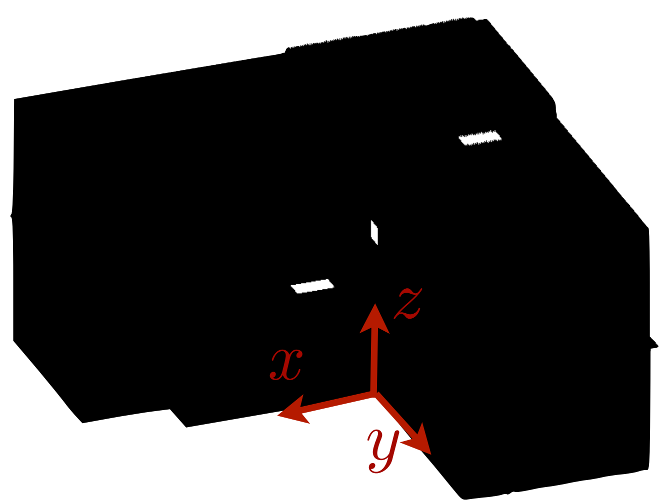




Oracle Dictionary Design

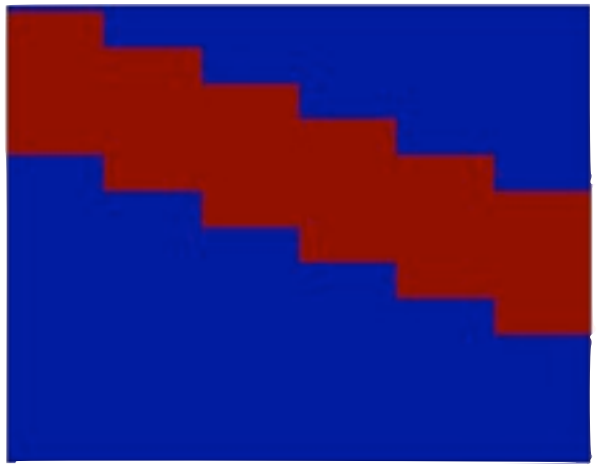
- A synthetic scene with a single depth object.
- The EPI consists of a line with slope relative to the object depth.

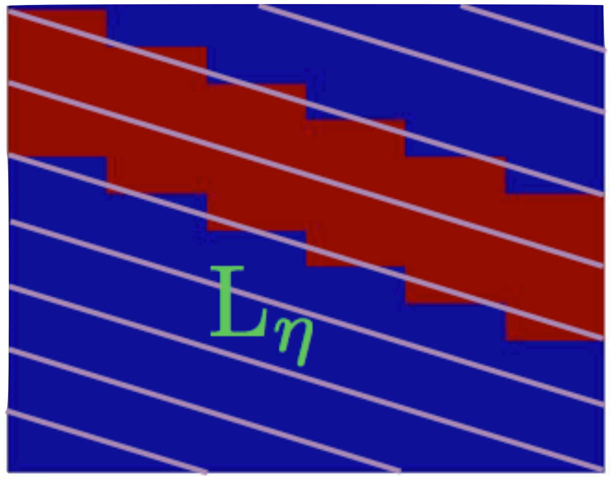


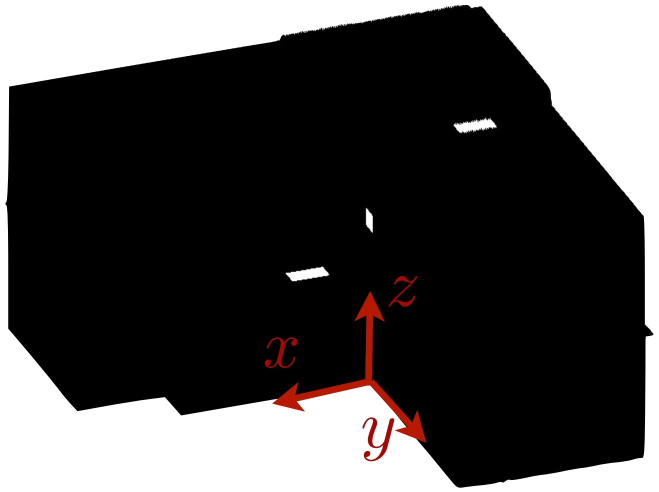


Oracle Dictionary Design

- A synthetic scene with a single depth object.
- The EPI consists of a line with slope relative to the object depth.
- Select a proper slope η for reordering lines L_η.

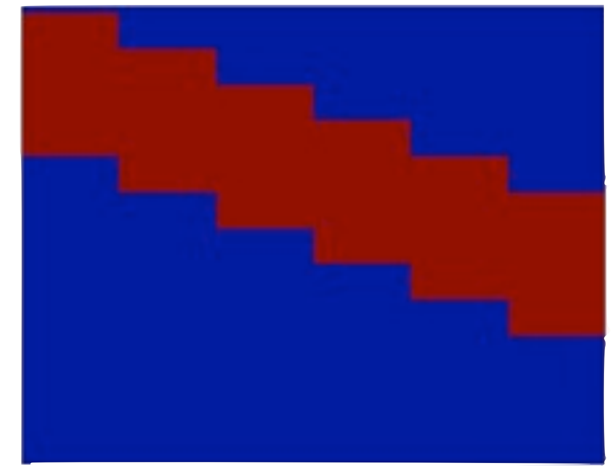


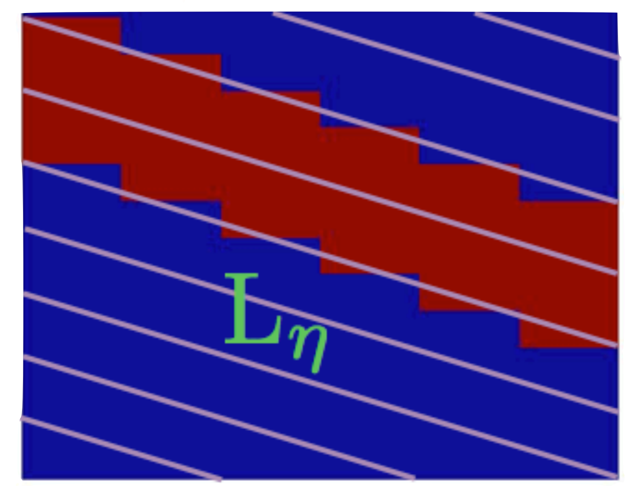


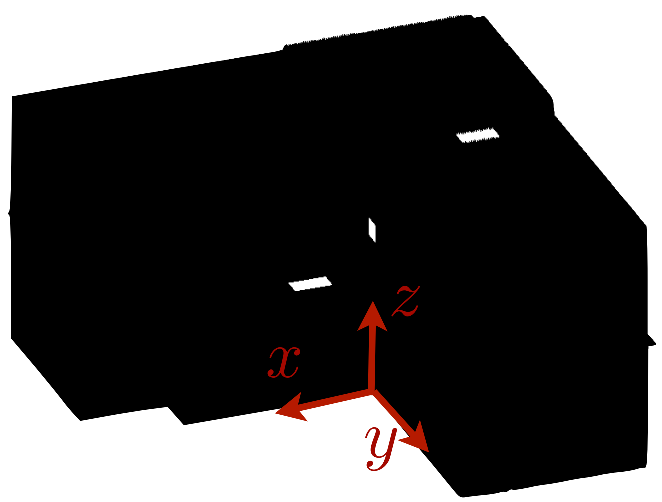


Oracle Dictionary Design

- A synthetic scene with a single depth object.
- The EPI consists of a line with slope relative to the object depth.
- Select a proper slope η for reordering lines L_η.
- The reordered image is piecewise constant.

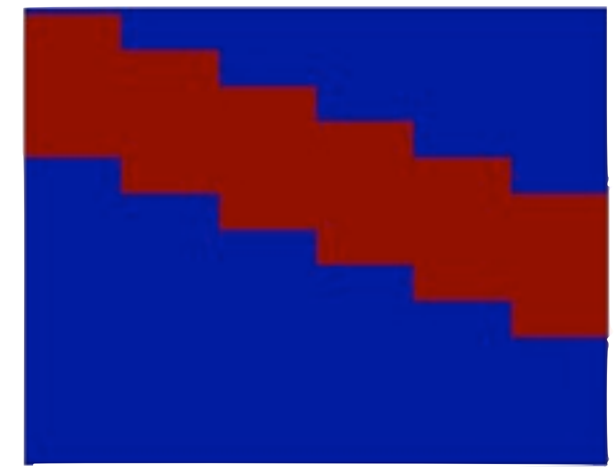


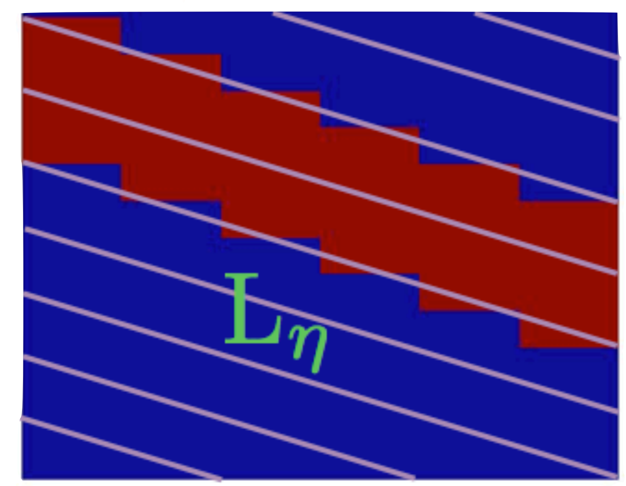


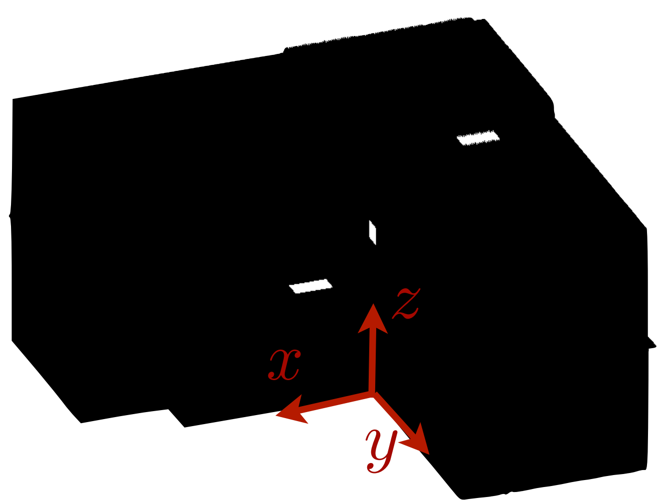


Oracle Dictionary Design

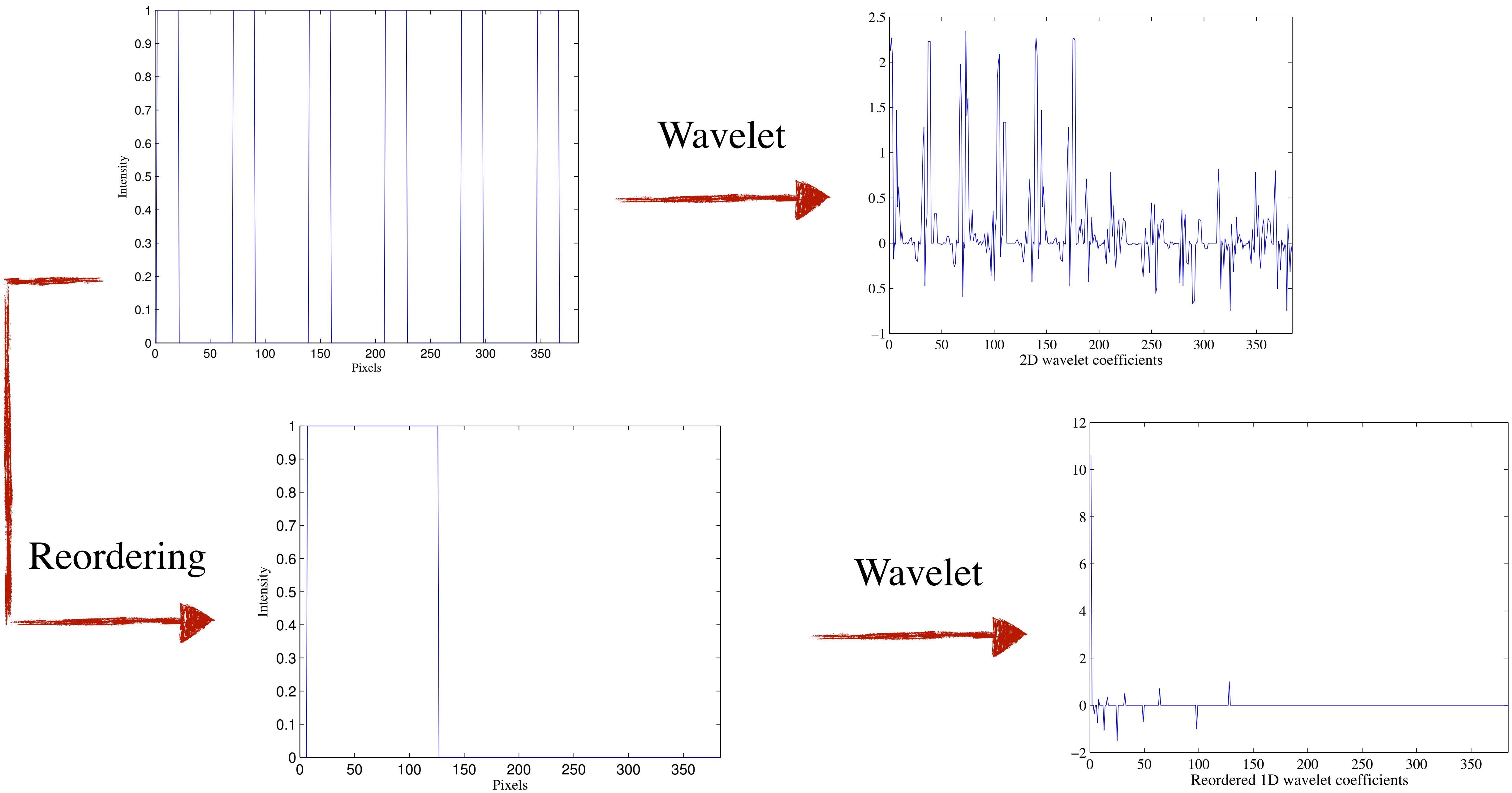
- A synthetic scene with a single depth object.
- The EPI consists of a line with slope relative to the object depth.
- Select a proper slope η for reordering lines L_η.
- The reordered image is piecewise constant.
  - Sparse in 1D wavelet transform





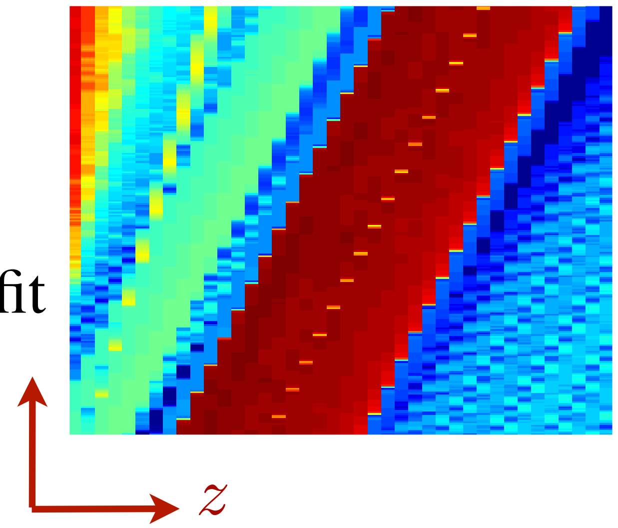


Dictionary Design



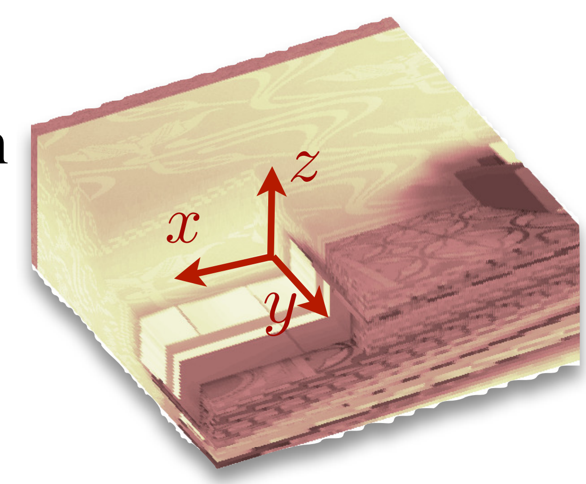
Dictionary Design

- Complex scenes consist of objects with different depth levels.
  - EPI lines do not have any preferential orientation.
- Using a redundant dictionary composed of union of bases to benefit from different reordering directions for EPIs.



$$\mathbf{\Psi} = \left[\, \mathbf{\Phi}_1^{\eta_1}, \mathbf{\Phi}_2^{\eta_2}, \cdots, \mathbf{\Phi}_\gamma^{\eta_\gamma} \, \right] \quad \mathbf{\Psi} \in \mathbb{R}^{ik \times \gamma ik}$$

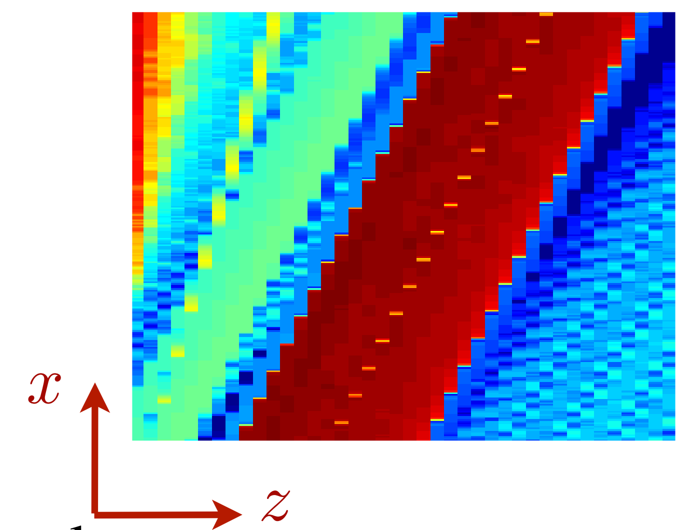
- Applying 1D wavelet transform along the remaining dimension on the image volume.



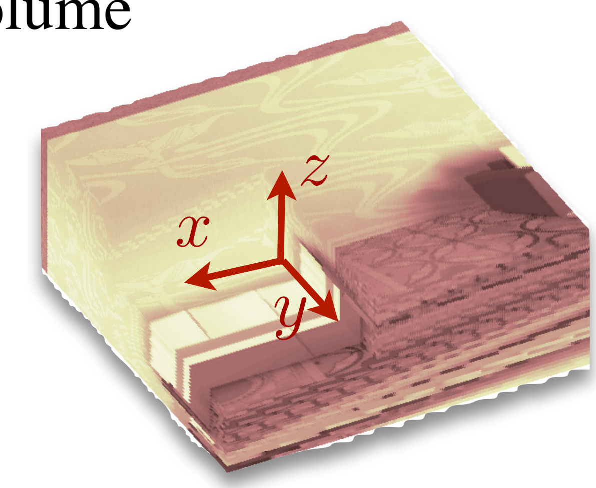
$$\mathbf{\Gamma} \in \mathbb{R}^{j \times j}$$

Dictionary Design

- The whole volume is sparse in the dictionary made of:
  1. The union of bases of 1D wavelet transforms on different reordering directions for EPIs.
  2. 1D wavelet transform along the 3rd direction of the volume.



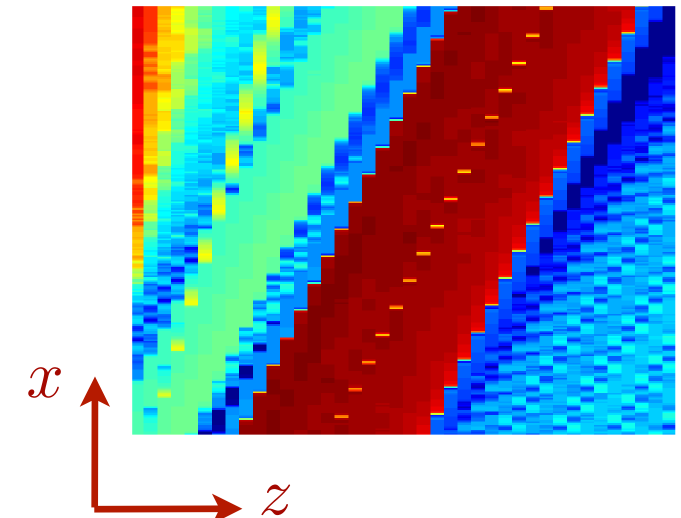
$\mathcal{X} \in \mathbb{R}^{i \times j \times k}$: image volume $\quad \widehat{\mathbf{X}} \in \mathbb{R}^{ik \times j}$: reshaped image volume



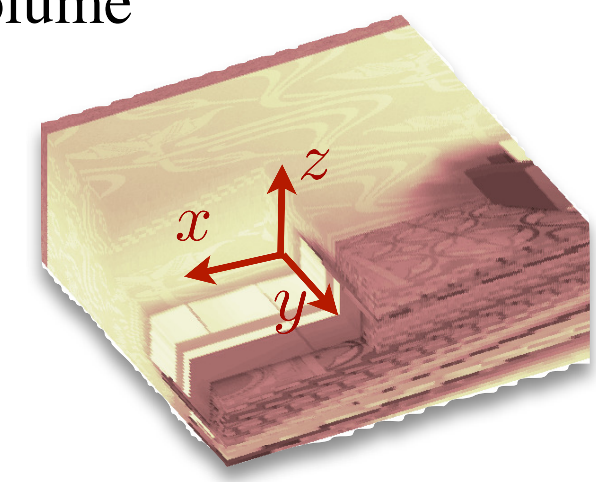
$\boldsymbol{\Theta} \in \mathbb{R}^{\gamma ik \times j}$: sparse matrix of coefficients

Dictionary Design

- The whole volume is sparse in the dictionary made of:
  1. The union of bases of 1D wavelet transforms on different reordering directions for EPIs.
  2. 1D wavelet transform along the 3rd direction of the volume.



$\mathcal{X} \in \mathbb{R}^{i \times j \times k}$: image volume $\widehat{\mathbf{X}} \in \mathbb{R}^{ik \times j}$: reshaped image volume

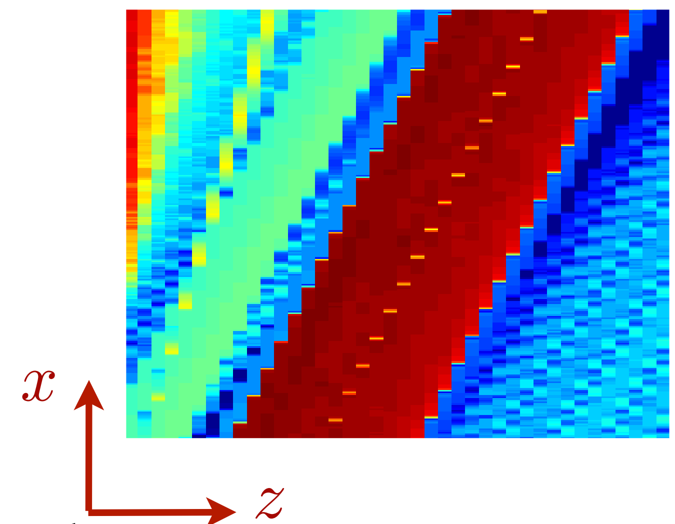


$\boldsymbol{\Theta} \in \mathbb{R}^{\gamma ik \times j}$: sparse matrix of coefficients

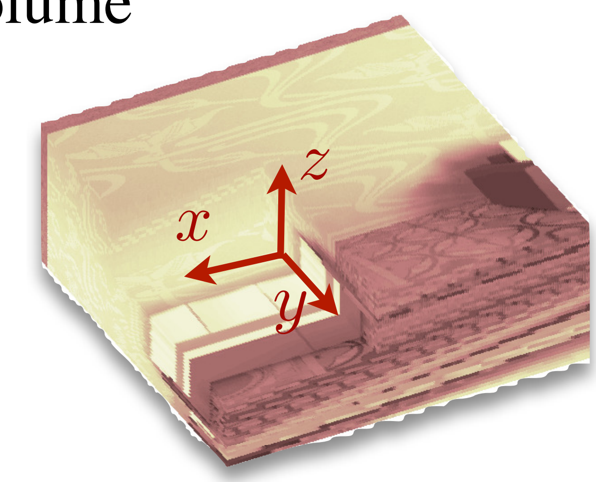
$$\widehat{\mathbf{X}} = \boldsymbol{\Psi}\boldsymbol{\Theta}\boldsymbol{\Gamma}^T$$

Dictionary Design

- The whole volume is sparse in the dictionary made of:
  1. The union of bases of 1D wavelet transforms on different reordering directions for EPIs.
  2. 1D wavelet transform along the 3rd direction of the volume.



$\mathcal{X} \in \mathbb{R}^{i \times j \times k}$: image volume $\widehat{\mathbf{X}} \in \mathbb{R}^{ik \times j}$: reshaped image volume

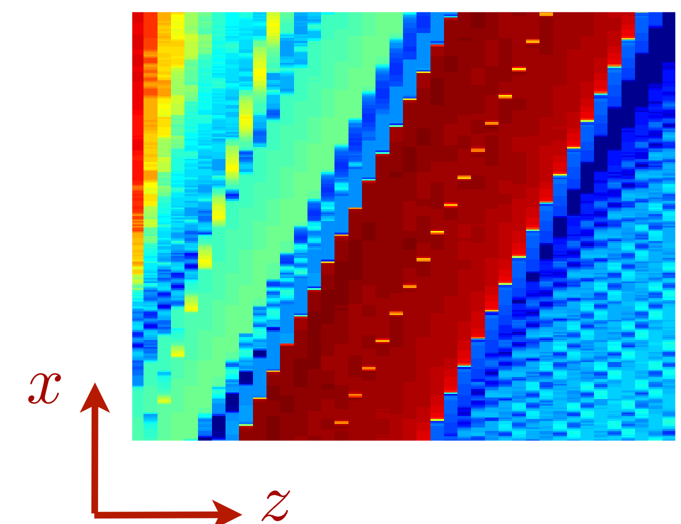


$\mathbf{\Theta} \in \mathbb{R}^{\gamma ik \times j}$: sparse matrix of coefficients

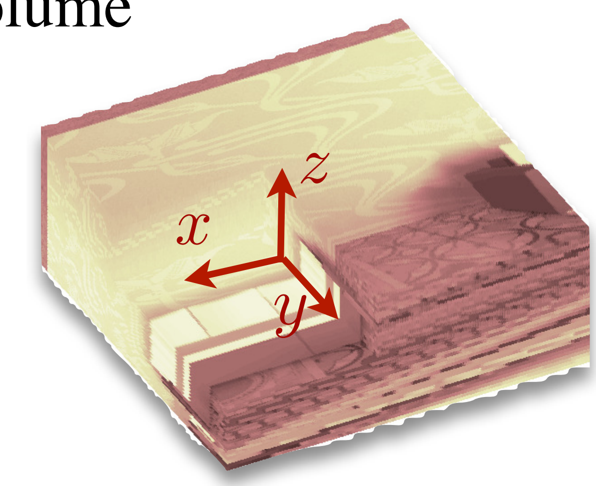
$$\widehat{\mathbf{X}} = \mathbf{\Psi \Theta \Gamma}^T \longrightarrow \widehat{X}_{vec} = \mathbf{\Omega} \Theta_{vec}$$

Dictionary Design

- The whole volume is sparse in the dictionary made of:
  1. The union of bases of 1D wavelet transforms on different reordering directions for EPIs.
  2. 1D wavelet transform along the 3rd direction of the volume.



$\mathcal{X} \in \mathbb{R}^{i \times j \times k}$: image volume $\qquad \widehat{\mathbf{X}} \in \mathbb{R}^{ik \times j}$: reshaped image volume



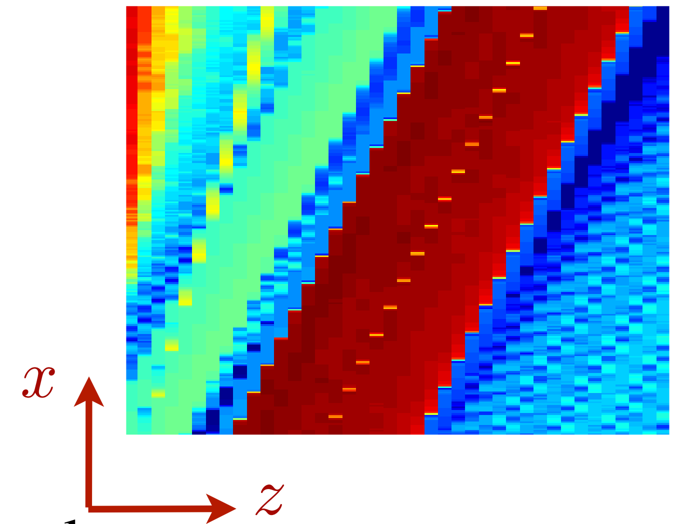
$\mathbf{\Theta} \in \mathbb{R}^{\gamma ik \times j}$: sparse matrix of coefficients

$$\widehat{\mathbf{X}} = \mathbf{\Psi \Theta \Gamma}^T \longrightarrow \widehat{X}_{vec} = \mathbf{\Omega} \Theta_{vec}$$

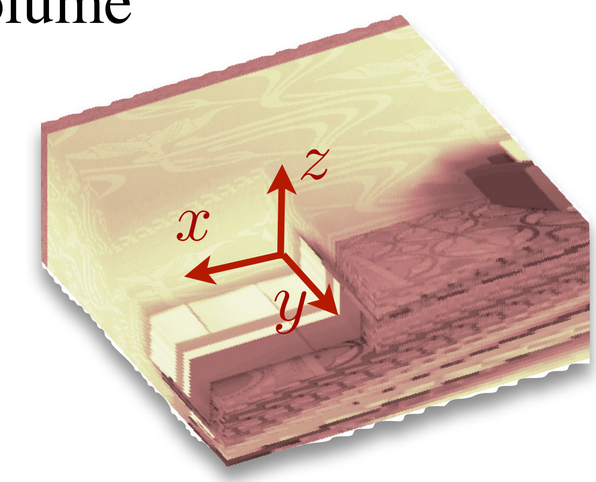
$$\mathbf{\Omega} = \mathbf{\Psi} \otimes \mathbf{\Gamma}$$

Dictionary Design

- The whole volume is sparse in the dictionary made of:
  1. The union of bases of 1D wavelet transforms on different reordering directions for EPIs.
  2. 1D wavelet transform along the 3rd direction of the volume.



$\mathcal{X} \in \mathbb{R}^{i \times j \times k}$: image volume $\quad \widehat{\mathbf{X}} \in \mathbb{R}^{ik \times j}$: reshaped image volume



$\mathbf{\Theta} \in \mathbb{R}^{\gamma ik \times j}$: sparse matrix of coefficients

$$\widehat{\mathbf{X}} = \mathbf{\Psi}\mathbf{\Theta}\mathbf{\Gamma}^T \longrightarrow \widehat{X}_{vec} = \mathbf{\Omega}\Theta_{vec}$$

$$\mathbf{\Omega} = \mathbf{\Psi} \otimes \mathbf{\Gamma} \longrightarrow \mathbf{\Omega} \in \mathbb{R}^{nk \times \gamma nk}$$

Measurement Scheme

- Random convolution measurement model for each camera.

k : number of cameras in the array

$\mathcal{A}_i \;\; 1 \leq i \leq k$: random convolution measurement matrix[1]

$x_i \;\; 1 \leq i \leq k$: image vector of each camera

$y_i \;\; 1 \leq i \leq k$: measurement vector for each camera

$$Y = \mathbf{A}X$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} \mathcal{A}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathcal{A}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathcal{A}_k \end{bmatrix}$$

[1] J. Romberg, "Compressive Sensing by Random Convolution", SIAM SIIMS, 09,

L. Jacques, et al., "CMOS Compressed Imaging by Random Convolution", ICASSP 09

Recovery Model

- Redundant Dictionary:

$$\underset{\Theta_{vec}\in\mathbf{R}^{\gamma nk}}{\text{argmin}} \ ||\Theta_{vec}||_1 \quad \text{subject to} \quad ||Y - \widehat{\mathbf{A}}\mathbf{\Omega}\Theta_{vec}||_2 \leq \epsilon$$

$$\widehat{X}_{vec} = \mathbf{\Omega}\Theta_{vec}$$

→ Joint Recovery

- 2D wavelet:

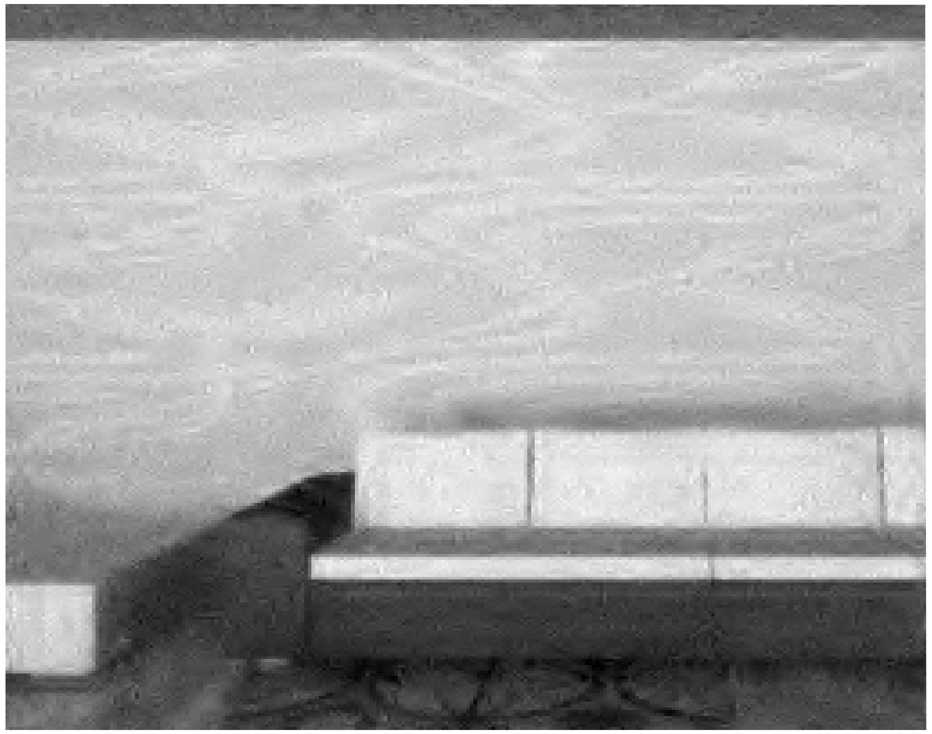
$$\underset{u_p\in\mathbb{R}^n}{\text{argmin}} \ ||u_p||_1 \quad \text{subject to} \quad ||y_p - \mathcal{A}_p\mathbf{\Phi}u_p||_2 \leq \epsilon_p$$

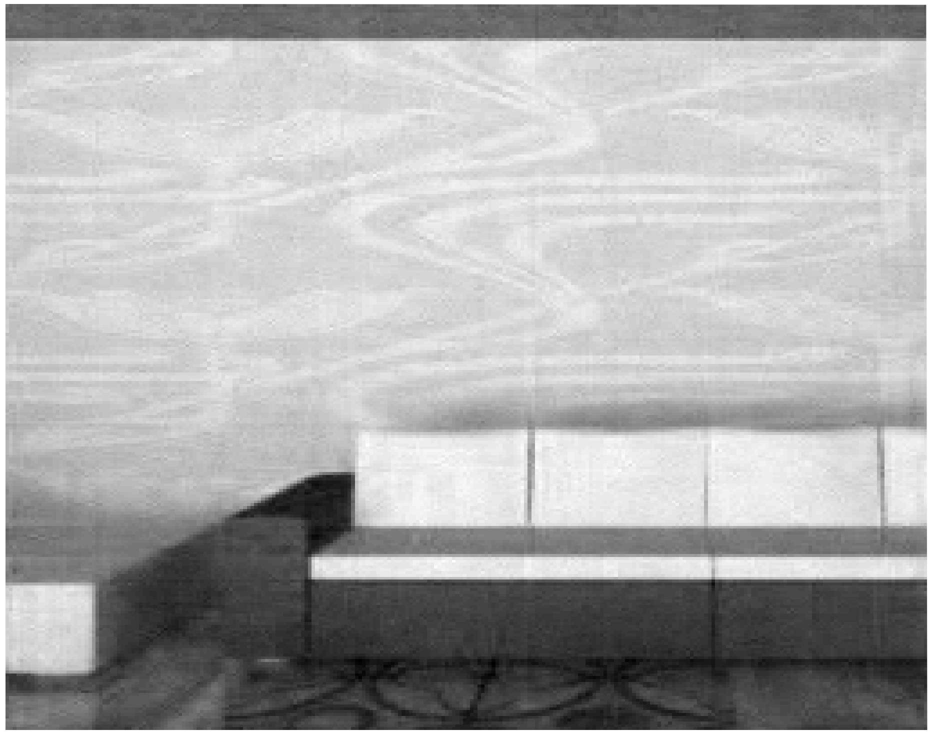
$$x_p = \mathbf{\Phi}u_p$$

→ Separate Recovery

Results

2D wavelet



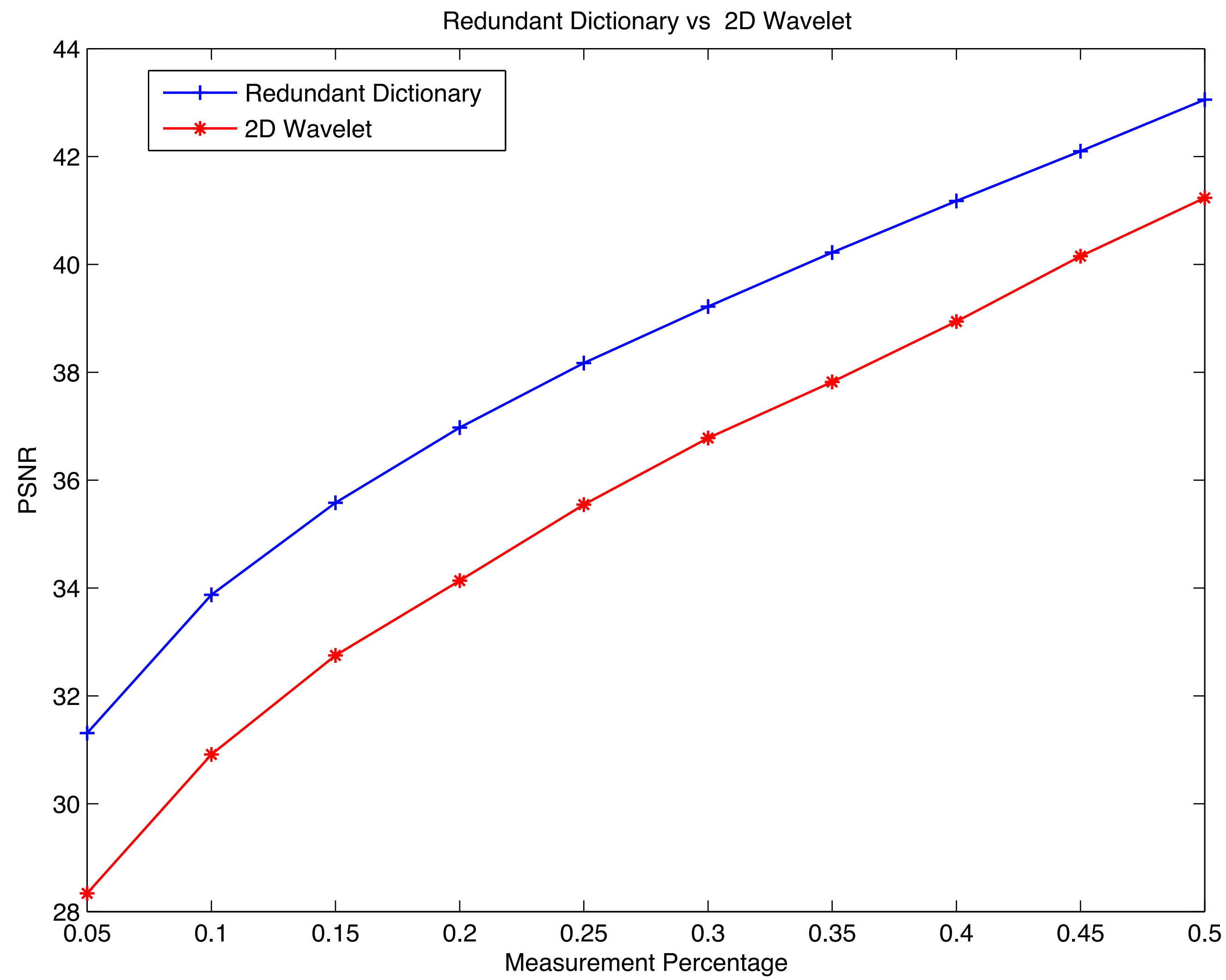


Redundant Dictionary

about 3dB improvement in average

sampling rate: 25% of image size

Results



Redundant Dictionary vs 2D Wavelet
Redundant Dictionary
2D Wavelet
PSNR
Measurement Percentage
44
42
40
38
36
34
32
30
28
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5

Recovery Model

- TV + Redundant Dictionary:

$$\underset{\Theta_{vec}\in\mathbf{R}^{\gamma nk}}{\text{argmin}} \; ||\mathbf{\Omega}\Theta_{vec}||_{\text{TV}} + \lambda||\Theta_{vec}||_1 \quad \text{subject to} \quad ||Y - \widehat{\mathbf{A}}\mathbf{\Omega}\Theta_{vec}||_2 \leq \epsilon$$

$$\widehat{X}_{vec} = \mathbf{\Omega}\Theta_{vec}$$

→ Separate-Joint Recovery

Proximal Splitting Methods [1]

- TV-denoising:

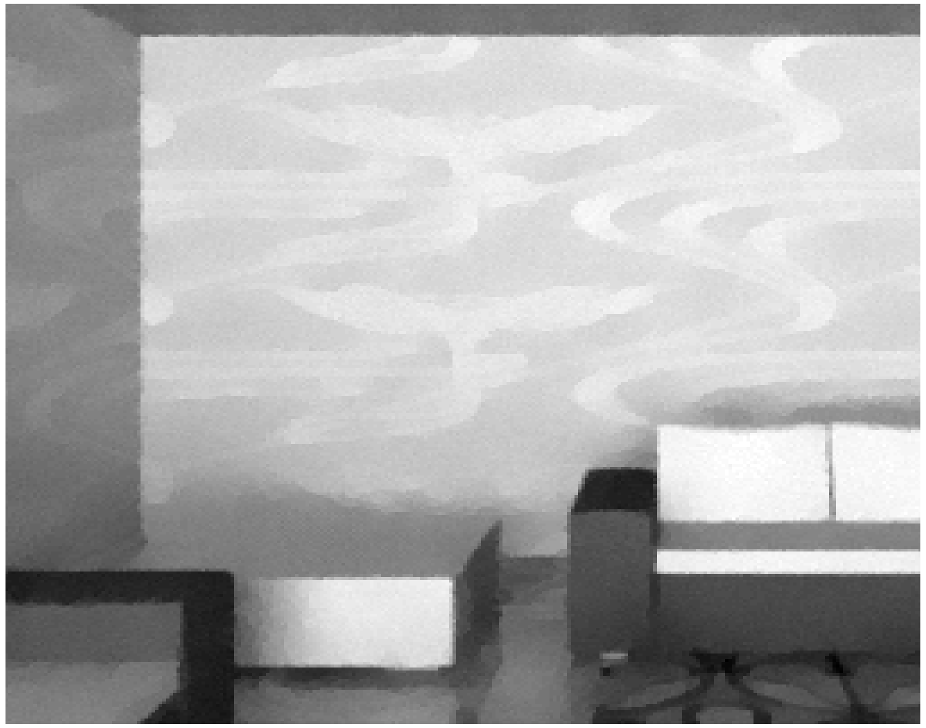
$$\underset{x_p\in\mathbf{R}^{n}}{\text{argmin}} \; ||x_p||_{\text{TV}} \quad \text{subject to} \quad ||y_p - \mathcal{A}_p x_p||_2 \leq \epsilon_p$$

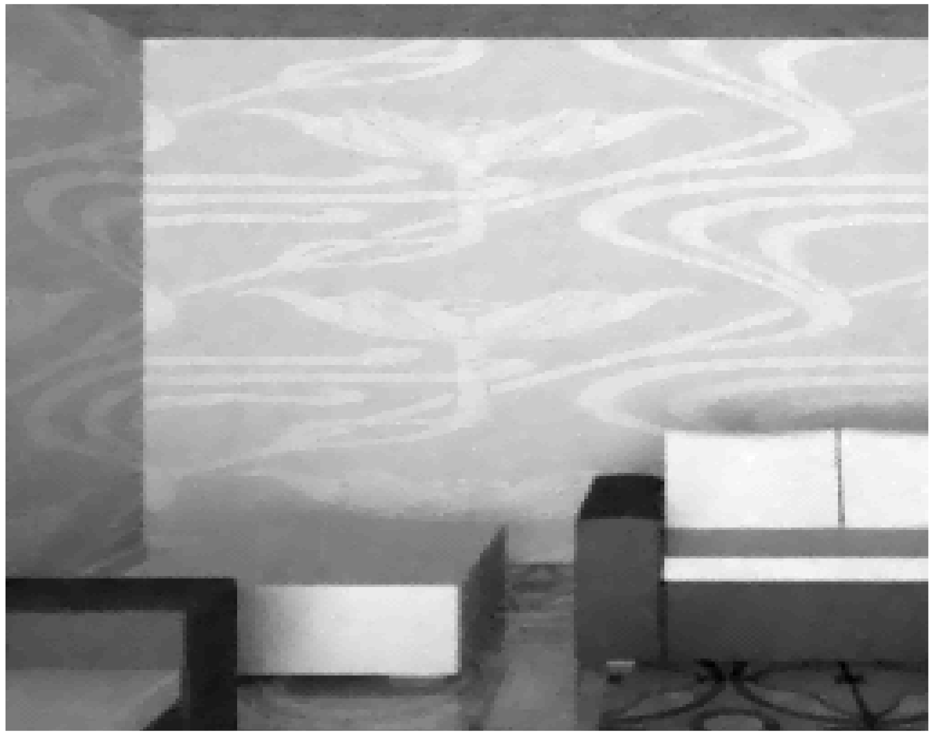
→ Separate Recovery

[1] P.L. Combettes et al., "Proximal Splitting Methods in Signal Processing", in Fixed-Point Alg. for Inv. Prob., 10

Results

TVDN



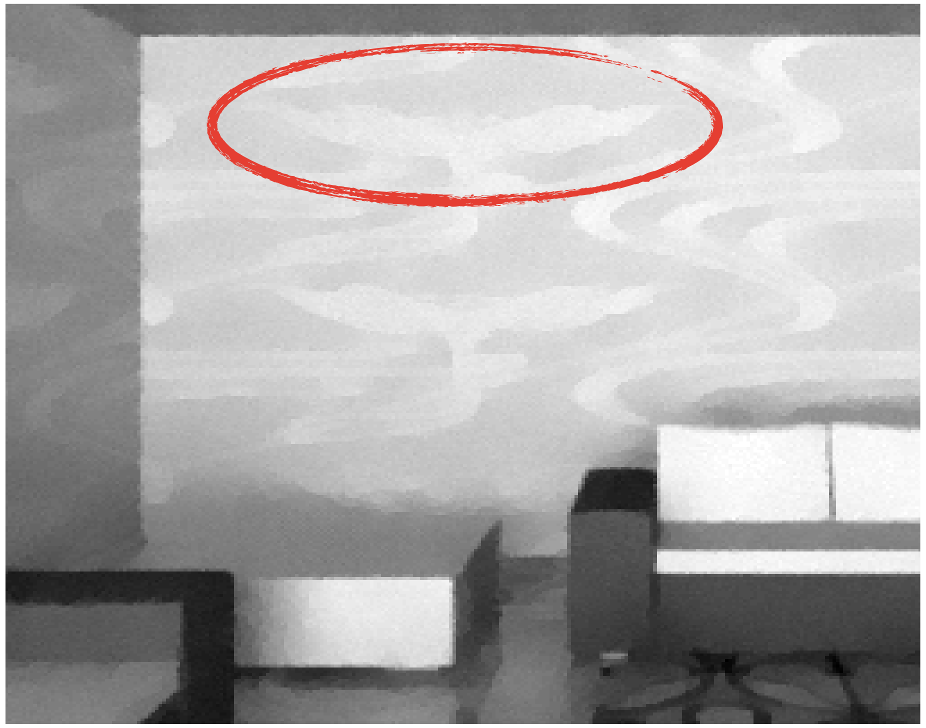


TV+Redundant Dictionary

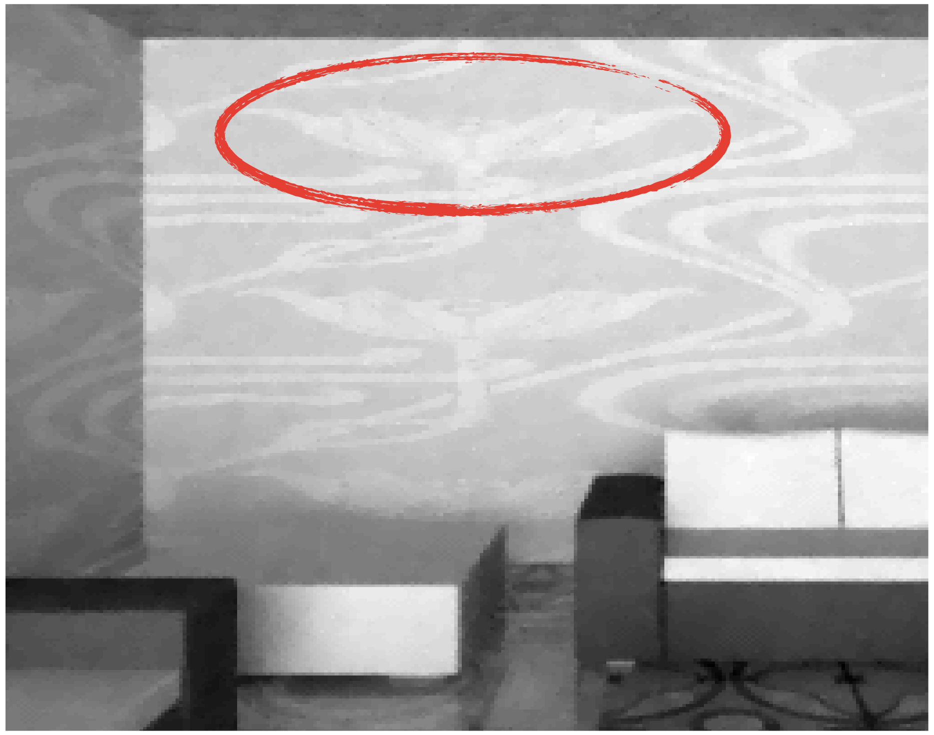
about 1.1 dB improvement in average

sampling rate: 25% of image size

Results



TVDN



TV+Redundant Dictionary

about 1.1 dB improvement in average

sampling rate: 25% of image size

Results

Original

TVDN

TV+Redundant Dictionary

Conclusion & Future Work

- Conclusion
  - Leveraging structure increases sparsity.
  - Reordered EPIs are sparse in 1D wavelet transform.
  - Recovery using TV and a redundant dictionary benefits from a separate and joint reconstruction scheme.
- Future Work
  - Designing a dictionary with less redundancy.
  - Considering joint low rank and sparsity scheme.

[1] J. Romberg, “Compressive Sensing by Random Convolution,” SIAM SIIMS, 2009.
[2] Edward H. Adelson and James R. Bergen, “The Plenop- tic Function and the Elements of Early Vision,” in Com- put. Mod. Vis. Proc., 1991.
[3] B.Wilburn,N.Joshi,V.Vaish,E.Talvala,E.Antnez,A.Barth, A. Adams, M. Horowitz, and M. Levoy, “High Performance Imaging using Large Camera Arrays,” ACM Trans. Graph., 2005.
[4] D. Donoho, “Compressed Sensing,” IEEE Trans. Inform. The- ory, 2006.
[5] E. Cande`s, J. Romberg, and T. Tao, “Stable Signal Recovery from Incomplete and Inaccurate Measurements,” Comm. Pure Appl. Math., 2005.
[6] E. Cande`s and J. Romberg and T. Tao, “Near Optimal Sig- nal Recovery From Random Projections: Universal Encoding Strategies?,” IEEE Trans. Inform. Theory, 2004.
[7] L. Jacques, P. Vandergheynst, A. Bibet, V. Majidzadeh, A. Schmid, and Y. Leblebici, “CMOS Compressed Imaging by Random Convolution,” ICASSP, 2009.
[8] D.Baron,M.F.Duarte,M.B.Wakin,S.Sarvotham,andR.G. Baraniuk, “Distributed Compressive Sensing,” CoRR, 2009.
[9] R. G. Baraniuk, “Compressive Sensing,” IEEE Sig. Proc. Mag., 2007.
[10] H. Rauhut, K. Schnass, and P. Vandergheynst, “Com- pressed Sensing and Redundant Dictionaries,” IEEE Trans. In- form. Theory, 2008.
[11] E. Cande`s and C. Eldar and D. Needell, “Compressed Sensing with Coherent and Redundant Dictionaries,” Appl. Comput. Harmon. Anal., 2011.
[12] G.Peyré andS.Mallat,“SurfaceCompressionwithGeometric Bandelets,” ACM SIGGRAPH, 2005.
[13] P. L. Combettes and J. C. Pesquet, “Proximal Splitting Meth- ods in Signal Processing,” arXiv/0912.3522, 2009.

Thank You