HLRTF: Hierarchical Low-Rank Tensor Factorization for Inverse Problems in Multi-Dimensional Imaging (CVPR 2022)

Yisi Luo[1], Xile Zhao[1], Deyu Meng[2,3], Taixiang Jiang[4]

[1]University of Electronic Science and Technology of China, [2]Xi'an Jiaotong University,
[3]Peng Cheng Laboratory, [4]Southwestern University of Finance and Economics
(contact: yisiluo1221@foxmail.com)

Section 1 (Classical Low-Tubal-Rank Model):

- Inverse problems in multi-dimensional imaging, e.g., completion, denoising, and compressive sensing, are challenging owing to the big volume of the data and the inherent ill-posedness.
- Recently, the tensor low-tubal-rank model was proposed for multi-dimensional image recovery. Tensor tubal rank is defined via the discrete Fourier transform (DFT):

$$\mathrm{rank}_t(\mathcal{A}) \triangleq \max_{i=1,2,\cdots,n_3} \{\mathrm{rank}((\mathcal{A}\times_3 \mathbf{F})^{(i)})\},$$

  where $\mathbf{F}$ denotes the DFT matrix. The tensor-tensor product (t-prod) is defined by: $\mathcal{A}*\mathcal{B} = ((\mathcal{A}\times_3\mathbf{F})\triangle(\mathcal{B}\times_3\mathbf{F}))\times_3\mathbf{F}^{-1}$ Based on the tubal-rank and t-prod, one can deduce the low-tubal-rank tensor factorization:

$$\mathrm{rank}_t(\mathcal{Y}*\mathcal{Z}) \leq \min\{\mathrm{rank}_t(\mathcal{Y}), \mathrm{rank}_t(\mathcal{Z})\}.$$

- Let $\mathcal{X} \in \mathbb{R}^{n_1\times n_2\times n_3}$ and $\mathrm{rank}_t(\mathcal{X}) = r$, then $\mathcal{X}$ can be factorized as $\mathcal{X} = \mathcal{A}*\mathcal{B}$, where $\mathcal{A} \in \mathbb{R}^{n_1\times r\times n_3}$ and $\mathcal{B} \in \mathbb{R}^{r\times n_2\times n_3}$.

Section 2 (Hierarchical Low-Rank Tensor Factorization):

- Considering the complex and diversified topology structures of real-world data, it is highly possible that the transform between the original tensor and the optimal low-rank representation is nonlinear and hierarchical, which can not be interpreted by the linear transform.
- Thus, we replace the linear DFT with a nonlinear deep neural network (DNN), which can obtain a better transformed low-rank representation:

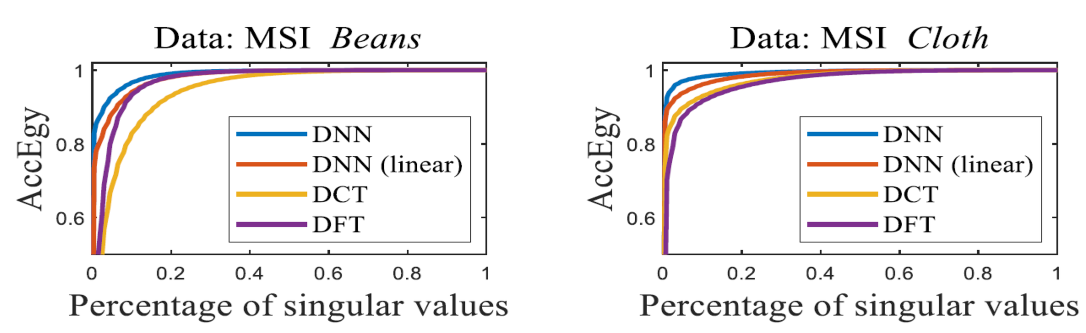


Fig 1: The AccEgy (AccEgy = $\sum_{i=1}^{k}\sigma_i^2/\sum_j\sigma_j^2$, where σ_i denotes the i-th singular value) w.r.t. the percentage of singular values of transformed frontal slices. DNN obtains a better low-rank representation than linear transforms.

- Based on the DNN transform, we define the hierarchical tubal-rank:

$$\mathrm{rank}_h(\mathcal{A}) \triangleq \max_{i=1,2,\cdots,n_3} \{\mathrm{rank}(f(\mathcal{A})^{(i)})\},$$

- where $f(\cdot)$ is a DNN. Its inverse DNN is denoted by $g(\cdot)$. The t-prod induced by $f(\cdot)$ is defined by $\mathcal{A}*_f\mathcal{B} = g(f(\mathcal{A})\triangle f(\mathcal{B}))$. We can deduce the corresponding hierarchical low-rank tensor factorization (HLRTF):

$$\mathrm{rank}_h(\mathcal{Y}*_f\mathcal{Z}) \leq \min\{\mathrm{rank}_h(\mathcal{Y}), \mathrm{rank}_h(\mathcal{Z})\}.$$

- Let $\mathcal{X} \in \mathbb{R}^{n_1\times n_2\times n_3}$ and $\mathrm{rank}_h(\mathcal{X}) = r$, then $\mathcal{X}$ can be factorized as $\mathcal{X} = \mathcal{A}*_f\mathcal{B}$, where $\mathcal{A} \in \mathbb{R}^{n_1\times r\times n_3}$ and $\mathcal{B} \in \mathbb{R}^{r\times n_2\times n_3}$.
- Based on the definition of t-prod, we can directly optimize the parameters of the inverse DNN $g(\cdot)$ and formulate the HLRTF model as

$$\min_{\hat{\mathcal{A}},\hat{\mathcal{B}},\{\mathbf{H}_j\}_{j=1}^{k}} L(g(\hat{\mathcal{A}}\triangle\hat{\mathcal{B}}),\mathcal{O}),$$

- where $\hat{\mathcal{A}} \triangleq f(\mathcal{A})$, $\hat{\mathcal{B}} \triangleq f(\mathcal{B})$, $\mathcal{O}$ is the observed data, and L is the fidelity loss. We use gradient descent to tackle this model.

Section 3 (Parametric Total Variation):

- In extreme slice missing cases, HLRTF suffers from vanishing gradient (see Fig. 2). Thus, we propose the parametric total variation (PTV) to constrain DNN parameters and tensor factor parameters:

$$\mathrm{PTV}(\Theta) \triangleq \|\nabla_x\hat{\mathcal{A}}\|_{\ell_1} + \|\nabla_y\hat{\mathcal{B}}\|_{\ell_1} + \|\nabla_x\mathbf{H}_k\|_{\ell_1}.$$

- where three terms $\|\nabla_x\hat{\mathcal{A}}\|_{\ell_1}$, $\|\nabla_y\hat{\mathcal{B}}\|_{\ell_1}$, and $\|\nabla_x\mathbf{H}_k\|_{\ell_1}$ respectively address the vanishing gradient in $\hat{\mathcal{A}}$, $\hat{\mathcal{B}}$, and $\mathbf{H}_k$.

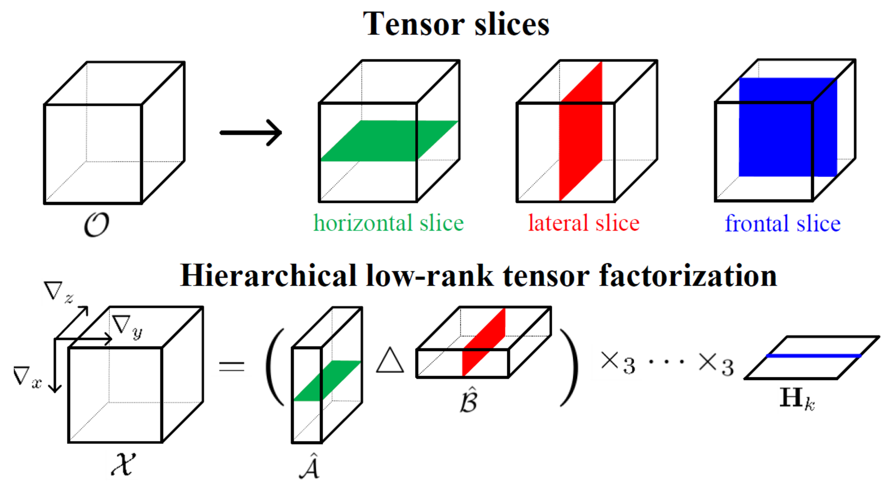
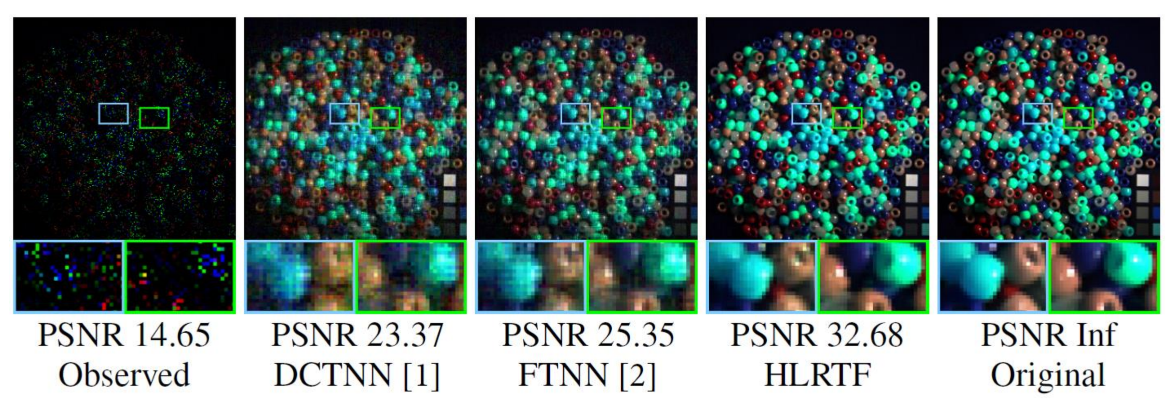


Fig 2: Illustrations of the vanishing gradient. (i) If the i-th horizontal slice of $\mathcal{O}$ is missed, then the gradient on the i-th horizontal slice of $\hat{\mathcal{A}}$ equals to zero. (ii) If the i-th lateral slice of $\mathcal{O}$ is missed, then the gradient on the i-th lateral slice of $\hat{\mathcal{B}}$ equals to zero. (iii) If the i-th frontal slice of $\mathcal{O}$ is missed, then the gradient on the i-th row of $\mathbf{H}_k$ equals to zero.

Section 4 (Experiments):

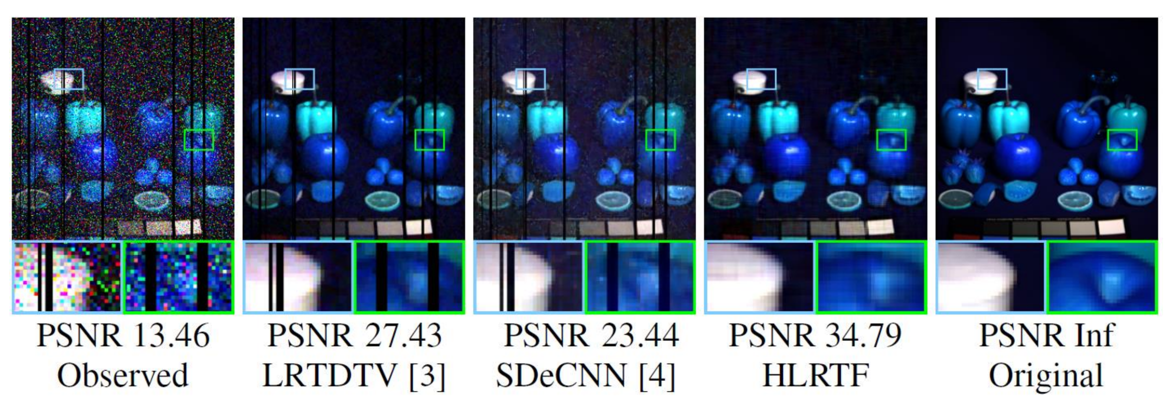
- By customizing different fidelity loss L, our method can be applied to different inverse problems in multi-dimensional imaging:
- (i) Multi-dimensional image inpainting ($L(\mathcal{X},\mathcal{O}) = \|(\mathcal{X}-\mathcal{O})_\Omega\|_F^2$)



[1] Low-rank tensor completion with a new tensor nuclear norm induced by invertible linear transforms, CVPR, 2019

[2] Framelet representation of tensor nuclear norm for third-order tensor completion, IEEE TIP, 2020

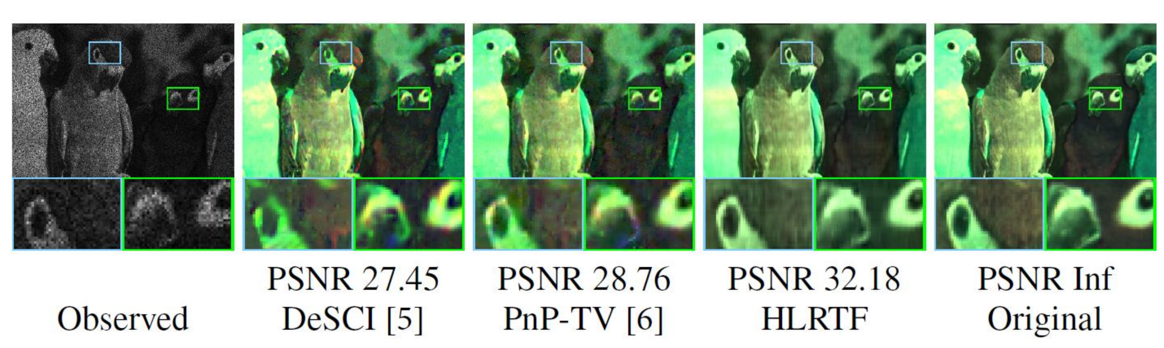
- (ii) Multispectral image denoising ($L(\mathcal{X},\mathcal{O}) = \|\mathcal{X}-\mathcal{O}\|_{\ell_1}$)



[3] Hyperspectral image restoration via total variation regularized low-rank tensor decomposition, IEEE JSTARS, 2017

[4] A single model CNN for hyperspectral image denoising, IEEE TGRS, 2020

- (iii) Snapshot compressive-spectral imaging ($L(\mathcal{X},\mathbf{O}) = \|\sum_{i=1}^{n_3}\mathcal{C}^{(i)}\odot\mathcal{X}^{(i)} - \mathbf{O}\|_F^2$)



[5] Rank minimization for snapshot compressive imaging, IEEE TPAMI, 2019

[6] Effective snapshot compressive-spectral imaging via deep denoising and total variation priors, CVPR, 2021

Contributions of this paper:

- We propose the HLRTF to capture the underlying low-rank structure of multi-dimensional images with compact representation abilities. We propose the PTV regularization to address the vanishing gradient issue.
- Extensive experiments validate the generalization abilities and effectiveness of HLRTF for different inverse problems in multi-dimensional imaging. *Code will be shared after request.*