Multi-dimensional Visual Data Completion via Low-rank Tensor Represention under Coupled Transform

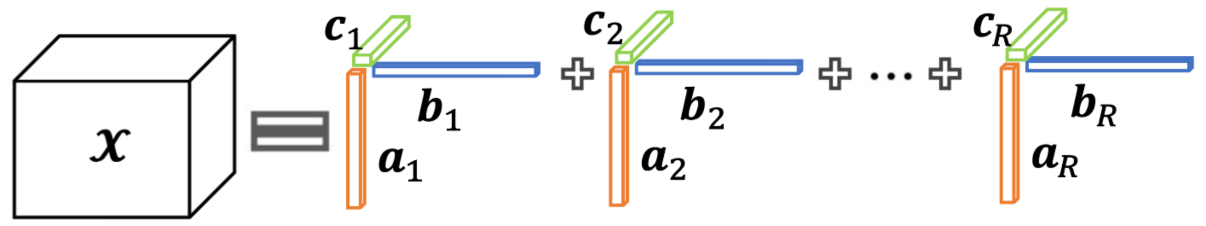
Jian-Li Wang[1], *Ting-Zhu Huang*[1,*], *Xi-Le Zhao*[1,*], *Tai-Xiang Jiang*[2], *Michael K. Ng*[3]

[1]University of Electronic Science and Technology of China; [2]Southwestern University of Finance and Economics; [3]University of Hong Kong

Background: Tensor Decomposition

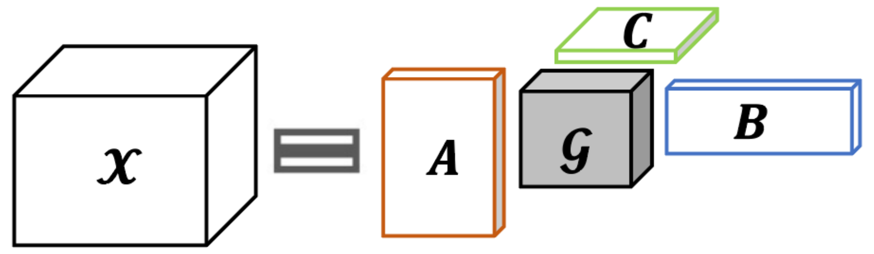
- Tensor decomposition aims to decompose a higher-order tensor to a set of low-dimensional factors and has powerful capability to capture the global correlations of data.
- CANDECOMP/PARAFAC (CP) decomposition:

$$\mathcal{X} = \sum_{r=1}^{R} \lambda_r a_r \circ b_r \circ c_r,$$



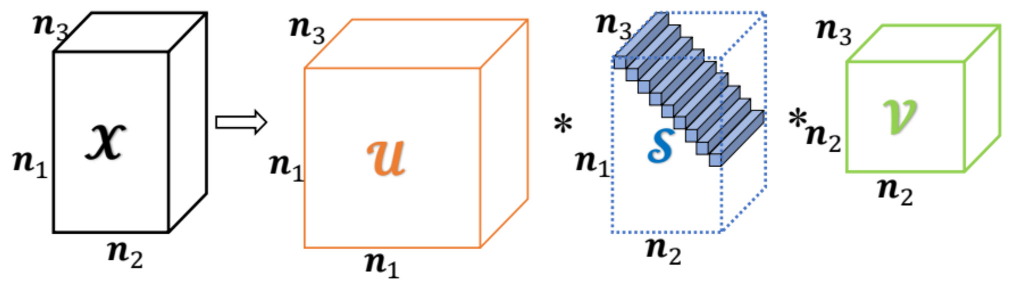
- Tucker decomposition:

$$\mathcal{X} = \mathcal{G} \times_1 A \times_2 B \times_3 C,$$



- Tensor singular values decomposition:

$$\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^{\mathbf{H}},$$



- They only consider the global data correlation (i.e., low-rankness) but ignore the spatial multi-scale nature.

Contributions

- We propose an enhanced low-rank tensor representation (LRTR) under coupled transform, which provides a novel perspective to exploit the implicit low-rank structure.
- We propose the CT-LRTC model by the enhanced LRTR for multidimensional visual data restoration.
- Extensive real examples on color images, multispectral images, and videos illustrate the proposed method outperms many atate-of-the-art methods in qualitative and quantitative aspects.

Enhanced Low-Rank Tensor Representation

- The enhanced LRTR under coupled transform aims to explore suitate transform to decorrelate the spatial and temporal/spectral dimensions, achieving a batter low-rank approximation.

- In the frist layer, we use a two-dimensional framelet transform to describe the local spatial correlation.

$$\mathcal{B}_1(\mathcal{X}) = \mathcal{X} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2,$$

where $\mathbf{W}_k$ (k = 1, 2) are the framelet transform matrix.

- In the second layer, we use a Fourier transform to characterize the global temporal/spectral correlation.

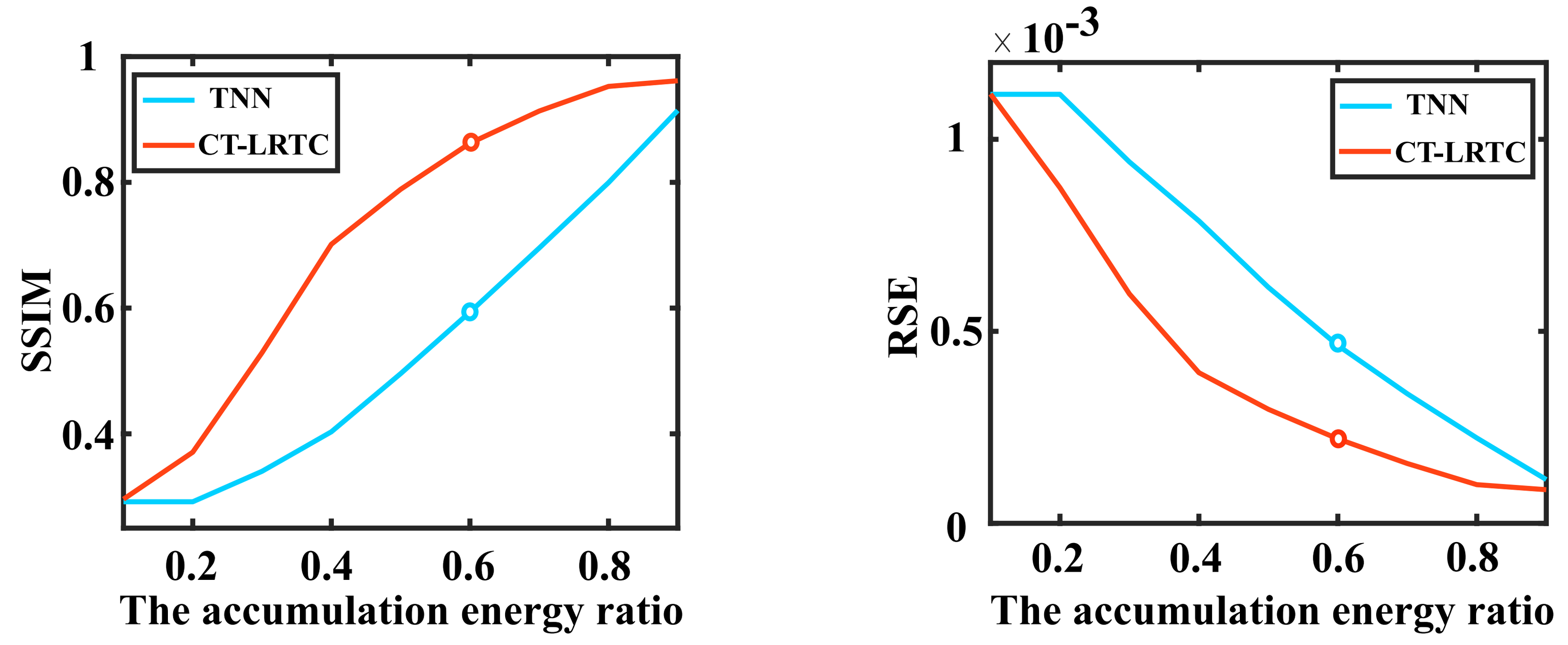
$$\mathcal{B}_2(\mathcal{B}_1(\mathcal{X})) = \mathcal{X} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{F},$$

where $\mathbf{F}$ the a Fourier transform matrix.

- In the third layer, we use a Karhunen-Loeve transform (via SVD) to characterize the global spatial correlation.

$$\mathcal{B}_3(\mathcal{B}_2(\mathcal{B}_1(\mathcal{X}))) = \mathbf{KL}(\mathcal{X} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{F}),$$

where $\mathbf{KL}(\cdot)$ is the sum of singular values of each frontal slice of $\bullet$.



Coupled Transforms-Based TC Model

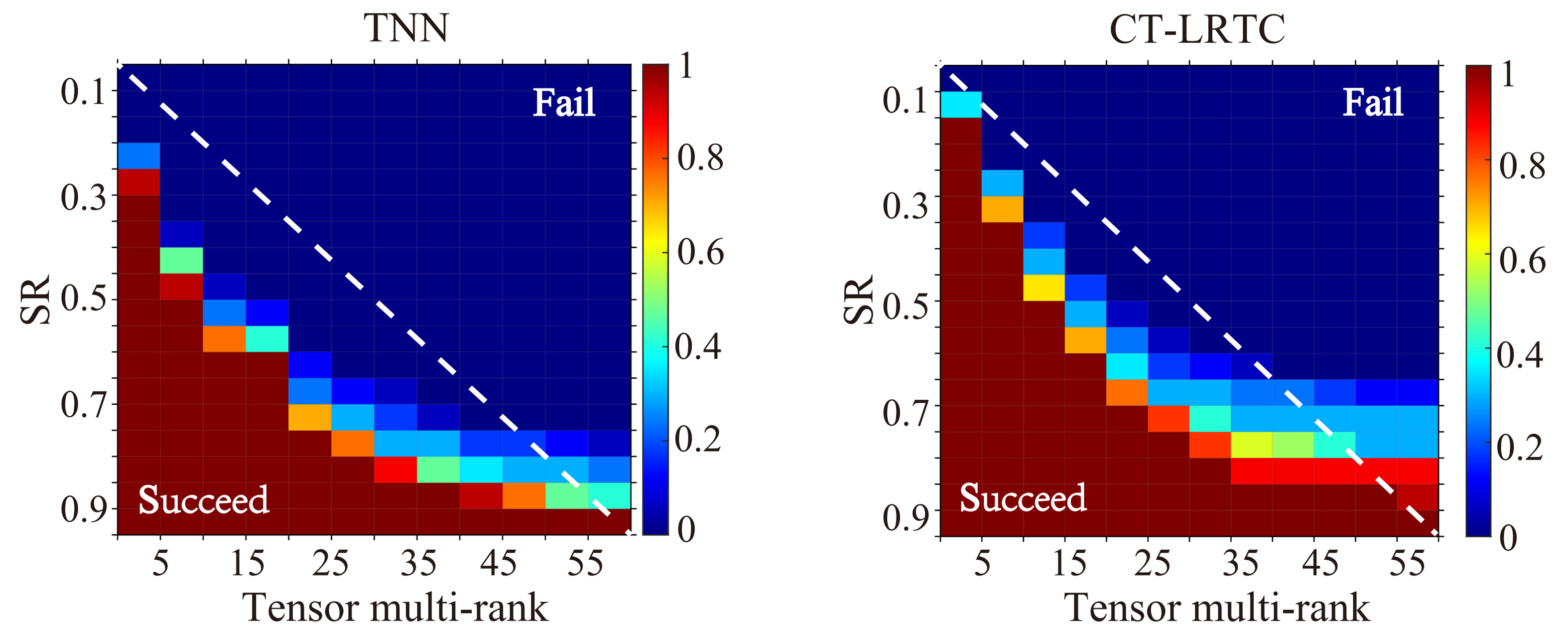
- Giving a partial observation $\mathcal{O}$ of the underlying tensor $\mathcal{X}$, the coupled transform-based tensor completion (CT-LRTC) model is:

$$\arg\min_{\mathcal{X}} \sum_{k=1}^{n_3} \left\| (\mathcal{X} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{F})^{(k)} \right\|_*$$
$$\text{s.t.} \quad \mathcal{P}_\Omega(\mathcal{X}) = \mathcal{P}_\Omega(\mathcal{O}).$$

Experimental Results

SR: sampling rate   PSNR: peak signal-to-noise ratio

- Synthetic Data Experiments (Success Rate)



- Multispectral Data Experiments (PSNR)

Dataset	SR	5%	10%	20%	Mean time (m)	Dataset	SR	5%	10%	20%	Mean time (m)
Clay	TNN	39.20	43.53	48.34	8.18	*Balloons*	TNN	34.99	39.61	44.84	10.43
	TNN-DCT	40.36	44.29	49.09	4.93		TNN-DCT	36.05	40.87	46.16	5.59
	t-TNN	37.80	42.58	48.73	5.12		t-TNN	38.37	43.21	48.30	4.79
	PSTNN	40.19	43.97	48.77	32.33		PSTNN	36.16	40.94	45.63	37.84
	TRLRF	41.86	43.75	47.79	58.78		TRLRF	36.43	40.13	42.35	60.84
	CT-LRTC	**44.71**	**48.16**	**51.92**	68.49		CT-LRTC	**41.47**	**45.07**	**48.99**	100.01

- Color Image Inpainting (Visual)