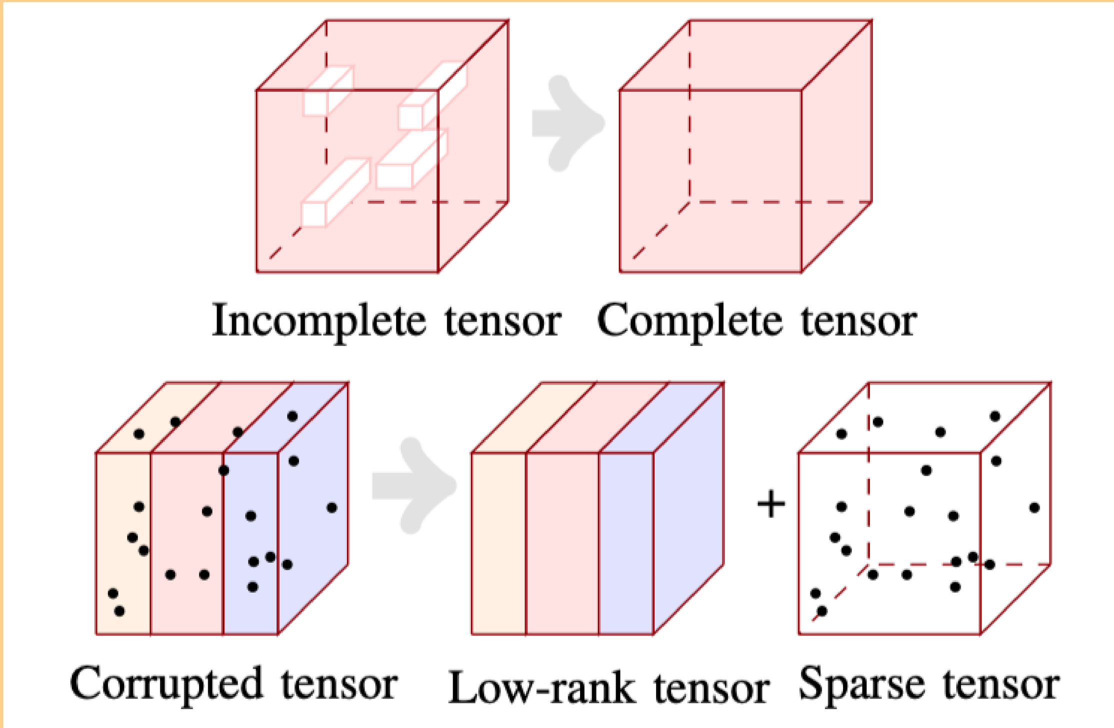
T-SVD Based Non-convex Tensor Completion and Robust Principal Component Analysis

Tao Li, Jinwen Ma

Abstract

Tensor completion and robust principal component analysis have been widely used in machine learning while the key problem relies on the minimization of a tensor rank that is very challenging. A common way to tackle this difficulty is to approximate the tensor rank with the ℓ_1-norm of singular values based on its Tensor Singular Value Decomposition (T-SVD). Besides, the sparsity of a tensor is also measured by its ℓ_1-norm. However, the ℓ_1 penalty is essentially biased and thus the result will deviate. In order to sidestep the bias, we propose a novel non-convex tensor rank surrogate function and a novel non-convex sparsity measure. In this new setting by using the concavity instead of the convexity, a majorization minimization algorithm is further designed for tensor completion and robust principal component analysis. Furthermore, we analyze its theoretical properties. Finally, the experiments on natural and hyperspectral images demonstrate the efficacy and efficiency of our proposed method.

Tensor Recovery



Perliminaries

Theorem 0.1 (T-SVD) *Suppose $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. Then there exists tensors $\mathcal{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}, \mathcal{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ and $\mathcal{S} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ such that $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^*$. Furthermore, $\mathcal{U}$ and $\mathcal{V}$ are orthogonal, while $\mathcal{S}$ is f-diagonal.*

Definition 0.1 (Tensor nuclear norm) *Let $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^*$ be the t-SVD of $\mathcal{A}$, the nuclear norm of $\mathcal{A}$ is defined as $\|\mathcal{A}\|_* = \sum_i \mathcal{S}(i,i,1) = \frac{1}{n_3}\sum_k \overline{\mathcal{S}}(i,i,k)$.*

Definition 0.2 (SCAD) *For some $\gamma > 2$ and $\lambda > 0$, the SCAD function is given by*

$$\varphi_{\lambda,\gamma}^{\mathrm{SCAD}}(t) = \begin{cases} \lambda|t| & \text{if} |t| \leq \lambda, \\ \frac{\gamma\lambda|t| - 0.5(t^2+\lambda^2)}{\gamma-1} & \text{if} \lambda < |t| < \gamma\lambda, \\ \frac{\gamma+1}{2}\lambda^2 & \text{if} |t| > \gamma\lambda. \end{cases}$$

Definition 0.3 (MCP) *For some $\gamma > 1$ and $\lambda > 0$, the MCP function is given by*

$$\varphi_{\lambda,\gamma}^{\mathrm{MCP}}(t) = \begin{cases} \lambda|t| - \frac{t^2}{2\gamma} & \text{if} |t| < \gamma\lambda, \\ \frac{\gamma\lambda^2}{2} & \text{if} |t| \geq \gamma\lambda. \end{cases}$$

Results

TABLE I
Tensor completion performances evaluation on natural images under varying sampling rates.

Method	20%			40%			60%			80%			time (s)
	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	PSNR	SSIM	FSIM	
SiLRTC	23.59	0.798	0.822	27.987	0.899	0.915	32.24	0.951	0.964	37.47	**0.977**	0.988	19.95
HaLRTC	23.82	0.797	0.828	28.39	0.902	0.920	33.038	0.953	0.968	39.27	**0.978**	0.991	31.32
FBCP	24.08	0.668	0.794	26.40	0.753	0.837	27.35	0.799	0.857	27.71	0.82	0.865	103.68
t-SVD	24.13	0.764	0.835	29.703	0.893	0.931	36.03	0.950	0.977	45.04	0.969	0.992	33.47
t-TNN	25.30	0.841	0.864	30.50	0.923	0.943	36.27	0.952	0.978	44.14	0.967	0.991	**3.03**
LRTC_{mcp}	**25.70**	**0.845**	**0.869**	**31.06**	**0.927**	**0.946**	**36.87**	**0.959**	**0.980**	**45.46**	0.973	**0.993**	3.79
LRTC_{scad}	**25.70**	**0.844**	**0.869**	**31.04**	**0.926**	**0.946**	**36.84**	**0.959**	**0.980**	**45.45**	0.973	**0.993**	3.83

TABLE II
Tensor completion performances evaluation on hyperspectral images under varying sampling rates. The unit is 10^{-4} for MSE.

Method	20%			40%			60%			80%			time (s)
	PSNR	MSE	ERGAS	PSNR	MSE	ERGAS	PSNR	MSE	ERGAS	PSNR	MSE	ERGAS	
SiLRTC	41.71	4.70	30.912	45.46	1.95	21.524	49.14	0.827	14.412	52.86	0.354	10.341	42.21
HaLRTC	42.11	4.53	29.626	45.95	1.90	20.556	49.79	0.801	13.514	53.67	0.342	9.660	53.11
FBCP	37.09	14.47	52.931	43.25	3.85	29.318	46.00	2.225	22.344	46.67	2.011	20.688	210.18
t-SVD	41.64	5.10	31.835	45.52	2.12	22.142	49.42	0.886	14.685	53.49	0.365	10.157	224.21
t-TNN	42.46	3.81	28.702	46.07	1.60	20.135	49.82	0.667	**13.272**	53.61	**0.290**	**9.515**	**47.29**
$\mathrm{LRTC}_{\mathrm{mcp}}$	**42.91**	**3.58**	**27.799**	**46.75**	**1.51**	**19.684**	**50.47**	**0.665**	**13.293**	**54.11**	**0.316**	**9.642**	70.95
$\mathrm{LRTC}_{\mathrm{scad}}$	**42.91**	**3.58**	**27.804**	**46.76**	**1.51**	**19.684**	**50.46**	**0.666**	13.298	**54.11**	**0.316**	**9.642**	71.14

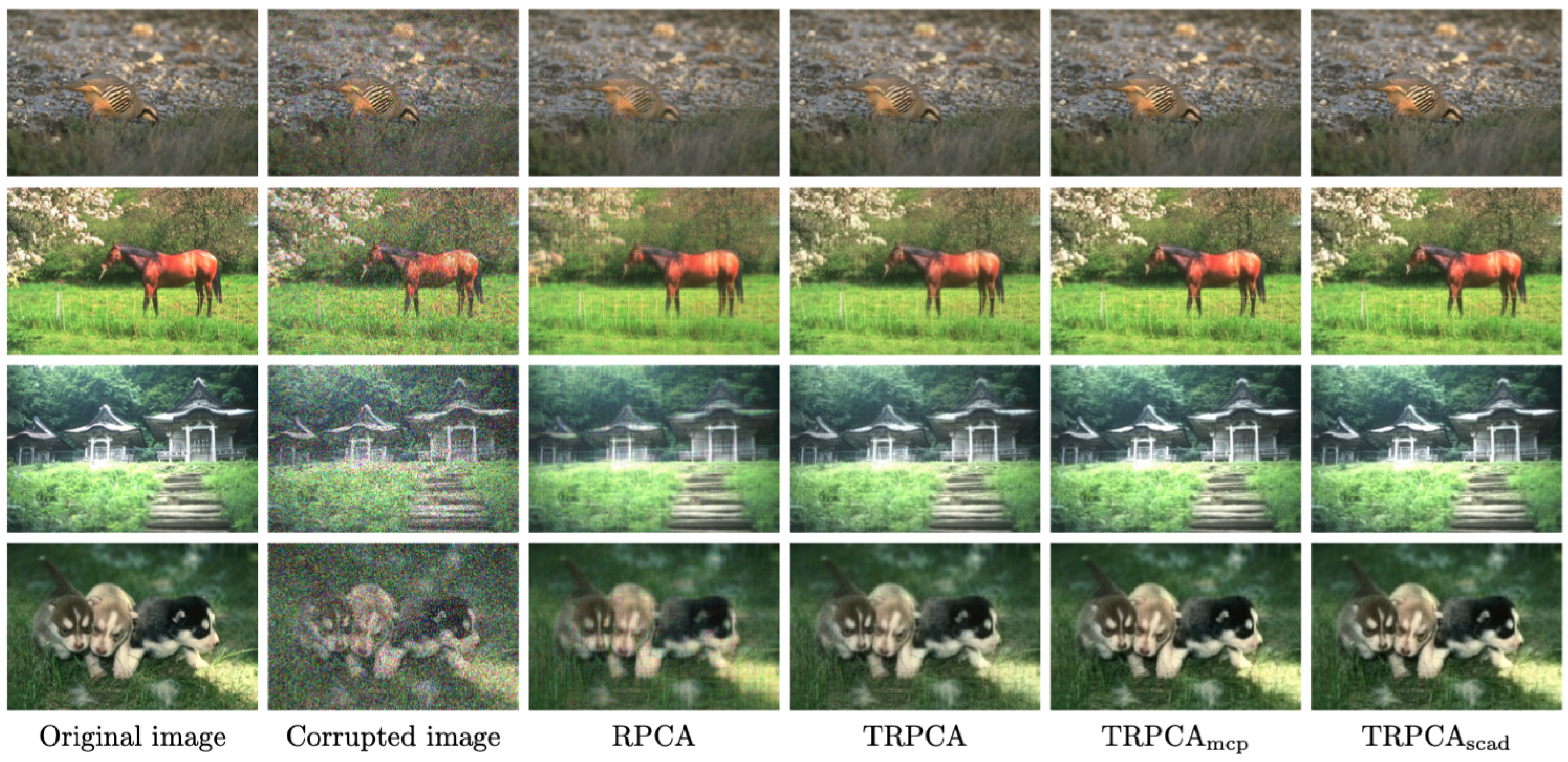


Fig. 4. Tensor RPCA performance comparison on example images. From top to bottom: $p_n = 0.1, 0.2, 0.3, 0.4$.

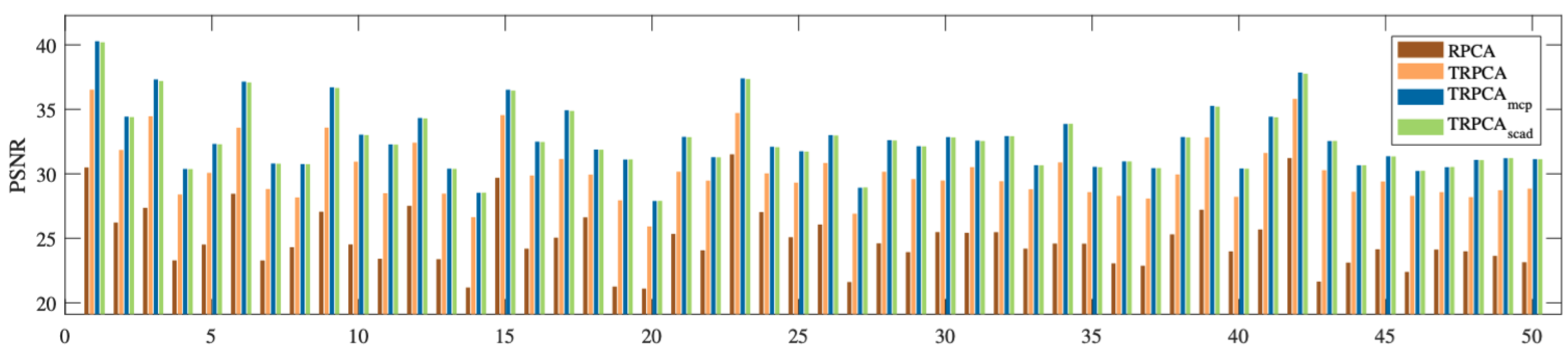


Fig. 5. Comparison of PSNR values obtained by RPCA, TRPCA, $\mathrm{TRPCA}_{\mathrm{mcp}}$,$\mathrm{TRPCA}_{\mathrm{scad}}$ on randomly selected 50 images.

Proposed

The novel tensor sparsity measure is defined as $\Phi_{\lambda,\gamma}(\mathcal{A}) = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\varphi_{\lambda,\gamma}(\mathcal{A}_{ijk})$.
Suppose $\mathcal{A}$ has t-SVD $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^*$, we define the $\gamma-$norm of $\mathcal{A}$ as $\|\mathcal{A}\|_\gamma = \frac{1}{n_3}\sum_{i,k}\varphi_{1,\gamma}(\overline{\mathcal{S}}(i,i,k))$.

Tensor Completion

Given a partially observed tensor $\mathcal{O} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$,Based on low rank assumption, tensor completion can be modeled as

$$\min_{\mathcal{X}} \operatorname{rank}(\mathcal{X}) \quad \text{s.t. } \mathcal{O}_\Omega = \mathcal{X}_\Omega.$$

Use the proposed tensor $\gamma-$norm to replace "rank" :

$$\min_{\mathcal{X}} \|\mathcal{X}\|_\gamma \quad \text{s.t. } \mathcal{O}_\Omega = \mathcal{X}_\Omega.$$

Majorization Minimization:

$$\min_{\mathcal{X}} Q_\gamma(\mathcal{X}|\mathcal{X}^{\mathrm{old}}) \quad \text{s.t. } \mathcal{O}_\Omega = \mathcal{X}_\Omega.$$

Tensor RPCA

Given a tensor $\mathcal{X}$, the goal of robust PCA is to decompose $\mathcal{X}$ into two parts: low-rank tensor $\mathcal{L}$ and sparse tensor $\mathcal{E}$. This problem can be formulated as

$$\min_{\mathcal{L},\mathcal{E}} \operatorname{rank}(\mathcal{L}) + \|\mathcal{E}\|_0 \quad \text{s.t. } \mathcal{L} + \mathcal{E} = \mathcal{X}.$$

Apply the proposed novel sparsity measure and tensor $\gamma-$norm, we obtain

$$\min_{\mathcal{L},\mathcal{E}} \|\mathcal{L}\|_{\gamma_1} + \Phi_{\lambda,\gamma_2}(\mathcal{E}) \quad \text{s.t. } \mathcal{L} + \mathcal{E} = \mathcal{X}.$$

Majorization minimization:

$$\min_{\mathcal{L},\mathcal{E}} Q_{\gamma_1}(\mathcal{L}|\mathcal{L}^{\mathrm{old}}) + Q_{\lambda,\gamma_2}(\mathcal{E}|\mathcal{E}^{\mathrm{old}}) \quad \text{s.t. } \mathcal{L} + \mathcal{E} = \mathcal{X}.$$