

引用格式:赵奉营,杨宏伟,赵丽娜. 增强的张量鲁棒主成分分析模型及其应用[J]. 北京化工大学学报(自然科学版), 2022, 49(4):105–116.

ZHAO FengYing, YANG HongWei, ZHAO LiNa. Applications of enhanced tensor robust principal component analysis[J]. Journal of Beijing University of Chemical Technology (Natural Science), 2022, 49(4):105–116.

# 增强的张量鲁棒主成分分析模型及其应用

赵奉营<sup>1</sup> 杨宏伟<sup>2</sup> 赵丽娜<sup>1\*</sup>

(北京化工大学 1. 数理学院; 2. 信息中心, 北京 100029)

**摘要:** 鲁棒主成分分析(RPCA)是处理图像恢复和背景建模问题的常用模型。针对原始RPCA及其改进模型对输入数据低秩结构的依赖性过强问题,提出一个增强的张量鲁棒主成分分析模型(E-TRPCA)并构造了一个新的增强张量核范数(E-TNN)正则项。E-TNN基于张量数据的低维子空间投影约束其低秩性,可以更真实地反映张量数据的潜在结构,增强模型的泛化性。利用交替方向乘子算法(ADMM)对目标函数进行优化求解,在图像去噪和背景建模上的实验结果表明所提方法在图像恢复效果和运行时间方面要优于当前的其他方法。

**关键词:** 张量鲁棒主成分分析; 低秩张量恢复; 增强张量核范数; 张量分解

**中图分类号:** O212 **DOI:** 10.13543/j.bhxbzr.2022.04.013

## 引言

鲁棒主成分分析(RPCA)是低秩恢复问题的一种。Candès等<sup>[1]</sup>提出基于矩阵的RPCA方法(matrix RPCA, MRPCA),能够从被稀疏大噪声损坏的观测数据中恢复出隐含的低秩分量;MRPCA将观测数据分解为低秩分量和稀疏分量,用矩阵核范数约束低秩分量,矩阵范数约束稀疏分量。该模型简单且易于求解,被广泛应用在视觉处理任务中,如图像恢复和背景建模<sup>[1–2]</sup>。但是MRPCA方法会将观测数据矩阵化,这样的预处理步骤会导致不必要的信息损失,带来次优的恢复结果。针对此问题,许多基于张量的RPCA方法被相继提出,如基于Tucker分解<sup>[3]</sup>的求和核范数(sum nuclear norm, SNN)模型<sup>[4]</sup>、基于张量火车分解<sup>[5]</sup>的张量火车核范数(tensor train nuclear norm, TTNN)模型<sup>[6]</sup>和基于张量环(tensor ring, TR)分解的鲁棒张量环补全(robust tensor ring completion, RTRC)模型<sup>[7]</sup>等。为了刻画张量不同模型之间的相关性,研究者们在

SNN、TTNN和RTRC中通过特殊的矩阵化策略将张量展开为一组矩阵,提出了可用交替方向乘子算法(ADMM)求解的张量鲁棒主成分分析(tensor-RPCA, TRPCA)模型。Kilmer等<sup>[8]</sup>提出了一种新的张量乘法 $t$ 积(tensor-tensor product)和张量奇异值分解(tensor singular value decomposition, t-SVD)策略。在 $t$ 积和t-SVD的张量框架下,Lu等<sup>[2]</sup>提出了张量核范数(tensor nuclear norm, TNN)概念,Zhou等<sup>[9]</sup>提出了张量低秩表示(low rank tensor representation, LRTR)模型。在TNN和LRTR模型中,以基于 $t$ 积的张量核范数t-TNN作为tubal秩的凸包络,t-TNN等价于张量的所有前切片组成的块循环矩阵的核范数或者张量按模3方向作离散傅里叶变换所得张量的所有前切片组成的块对角矩阵的核范数,因此t-TNN可以看作一种特殊的矩阵化方式,用来刻画张量的空间信息和第三通道的相关性。此外,t-SVD可以利用快速傅里叶变换加速计算,并且tubal秩可以描述张量子空间结构。上述模型在图像恢复和背景建模中取得了较好的效果,但是对输入数据低秩结构的强依赖性问题仍然没有得到解决。

针对上述问题,本文在 $t$ 积和t-SVD的张量框架下提出一个增强的张量鲁棒主成分分析模型(E-TRPCA)用于图像恢复和背景建模。首先学习得到一个字典张量,并构造了一个增强的张量核范

收稿日期: 2021-07-05

基金项目: 中央高校基本科研业务费专项资金

第一作者: 男, 1993年生, 硕士生

\* 通信联系人

E-mail: zhaoln@mail.buct.edu.cn

数(E-TNN)正则项;其次,在低秩张量表示的基础上,提出了一个去随机噪声的模型;最后,设计了一个高效的交替方向乘子算法来解决所提出的问题。在图像恢复和背景建模上的实验证明了所提方法的优越性。

## 1 张量奇异值分解框架

### 1.1 符号说明

给定张量  $\mathcal{A} \in \mathbb{R}^{h \times w \times s}$ ,  $\mathcal{A}_{(1, :, :)} \setminus \mathcal{A}_{(:, 1, :)}$  以及  $\mathcal{A}_{(:, :, 1)}$  分别是  $\mathcal{A}$  的第一个横切片、第一个侧切片和第一个前切片,  $\mathcal{A}_{i,j,k}$  是  $\mathcal{A}$  在  $(i, j, k)$  位置的元素。 $\text{unfold}()$  运算将张量展开为矩阵,  $\text{fold}()$  运算是其逆运算。

$$\text{unfold}(\mathcal{A}) = \begin{bmatrix} \mathcal{A}_{(:, :, 1)} \\ \mathcal{A}_{(:, :, 2)} \\ \vdots \\ \mathcal{A}_{(:, :, s)} \end{bmatrix}, \text{fold} \left( \begin{bmatrix} \mathcal{A}_{(:, :, 1)} \\ \mathcal{A}_{(:, :, 2)} \\ \vdots \\ \mathcal{A}_{(:, :, s)} \end{bmatrix} \right) = \mathcal{A}$$

$\overline{\mathcal{A}} = \text{fft}(\mathcal{A}, [], 3)$  记为对  $\mathcal{A}$  的每个管道的离散傅里叶变换。

$\text{bdiag}()$  运算将张量  $\overline{\mathcal{A}}$  变为块对角矩阵  $\overline{\mathbf{A}}$

$$\overline{\mathbf{A}} = \text{bdiag}(\overline{\mathcal{A}}) = \begin{bmatrix} \overline{\mathcal{A}}_{(:, :, 1)} & & & \\ & \overline{\mathcal{A}}_{(:, :, 2)} & & \\ & & \ddots & \\ & & & \overline{\mathcal{A}}_{(:, :, s)} \end{bmatrix}$$

$\text{bcirc}()$  运算是将张量  $\mathcal{A}$  变为块循环矩阵的运算

$$\text{bcirc}(\mathcal{A}) = \begin{bmatrix} \mathcal{A}_{(:, :, 1)} & \mathcal{A}_{(:, :, s)} & \cdots & \mathcal{A}_{(:, :, 2)} \\ \mathcal{A}_{(:, :, 2)} & \mathcal{A}_{(:, :, 1)} & \cdots & \mathcal{A}_{(:, :, 3)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{A}_{(:, :, s)} & \mathcal{A}_{(:, :, s-1)} & \cdots & \mathcal{A}_{(:, :, 1)} \end{bmatrix}$$

### 1.2 t 积和张量奇异值分解

**定义 1 (t 积)** 给定张量  $\mathcal{A} \in \mathbb{R}^{h \times l \times s}$ ,  $\mathcal{B} \in \mathbb{R}^{l \times w \times s}$ , 则 t 积定义为<sup>[8]</sup>

$$\mathcal{A} * \mathcal{B} = \text{fold}(\text{bcirc}(\mathcal{A}) \text{unfold}(\mathcal{B}))$$

**定义 2 (正交张量)** 如果  $\mathcal{Q} \in \mathbb{R}^{n \times n \times l}$  满足  $\mathcal{Q}^T * \mathcal{Q} = \mathcal{Q} * \mathcal{Q}^T = \mathcal{I} \in \mathbb{R}^{n \times n \times l}$ , 则  $\mathcal{Q}$  为正交张量。进一步, 如果  $\mathcal{Q} \in \mathbb{R}^{p \times q \times l}$  满足  $\mathcal{Q}^T * \mathcal{Q} = \mathcal{I} \in \mathbb{R}^{q \times q \times l}$ , 则  $\mathcal{Q}$  为部分正交张量。

**定义 3 (张量奇异值分解)**<sup>[2]</sup> 给定张量  $\mathcal{A} \in \mathbb{R}^{h \times w \times s}$ , 它可以分解为如下形式

$$\mathcal{A} = \mathcal{U} * \mathcal{I} * \mathcal{V}^T$$

式中,  $\mathcal{U} \in \mathbb{R}^{h \times h \times s}$ ,  $\mathcal{V} \in \mathbb{R}^{w \times w \times s}$  是正交张量,  $\mathcal{I} \in \mathbb{R}^{h \times w \times s}$  是  $f$ -diagonal 张量。

**定义 4 (管道秩)** 给定张量  $\mathcal{A} \in \mathbb{R}^{h \times w \times s}$ , 则张量的管道秩为  $\mathcal{I}$  中非零管道的个数<sup>[10]</sup>, 记为  $\text{rank}_t(\mathcal{A})$

$$\text{rank}_t(\mathcal{A}) = \{i, \mathcal{I}(i, i, :) \neq 0\}$$

式中,  $\mathcal{A} = \mathcal{U} * \mathcal{I} * \mathcal{V}^T$  是  $\mathcal{A}$  的 t-SVD 分解。

**定义 5 (张量核范数)**<sup>[2]</sup>  $\mathcal{A} = \mathcal{U} * \mathcal{I} * \mathcal{V}^T$  是  $\mathcal{A}$  的 t-SVD,  $\mathcal{A} \in \mathbb{R}^{h \times w \times s}$ , 则  $\mathcal{A}$  的核范数表示为

$$\|\mathcal{A}\|_* = \sum_{i=1}^r \mathcal{I}_{(:, :, i)}$$

其中,  $r = \text{rank}_t(\mathcal{A})$ 。

**引理 1** 张量  $\mathcal{A} \in \mathbb{R}^{h \times w \times s}$ ,  $\mathcal{B} \in \mathbb{R}^{w \times h \times s}$ , 令  $\mathcal{F} = \mathcal{A} * \mathcal{B}$ , 则有:

$$\textcircled{1} \quad \|\mathcal{A}\|_F^2 = \frac{1}{s} \|\overline{\mathbf{A}}\|_F^2, \text{ 并且 } \langle \mathcal{A}, \mathcal{B} \rangle = \frac{1}{s} \langle \overline{\mathbf{A}}, \overline{\mathbf{B}} \rangle;$$

$$\textcircled{2} \quad \mathcal{F} = \mathcal{A} * \mathcal{B}, \text{ 则 } \overline{\mathbf{F}} = \overline{\mathbf{A}} \overline{\mathbf{B}}.$$

**引理 2**<sup>[11]</sup>  $\forall \mathbf{A} \in \mathbb{R}^{m \times n}$ , 则问题

$$\min_{\mathbf{V} \mathbf{V}^T = \mathbf{I}} \langle \mathbf{A}, \mathbf{V} \rangle \quad (1)$$

有全局解  $\mathbf{V}^* = \mathbf{B} \mathbf{C}^T$ ,  $\mathbf{A} = \mathbf{B} \mathbf{C}^T$  是  $\mathbf{A}$  的奇异值分解。

根据引理 1, 本文给出了引理 2 的张量推广形式。

**定理 1** 任给张量  $\mathcal{A} \in \mathbb{R}^{h \times w \times s}$ , 问题

$$\min_{\mathcal{V} * \mathcal{V}^T = \mathcal{I}} \langle \mathcal{A}, \mathcal{V} \rangle \quad (2)$$

的全局解为  $\mathcal{V} = \mathcal{B} * \mathcal{D}^T$ ,  $\mathcal{A} = \mathcal{B} * \mathcal{C} * \mathcal{D}^T$  是  $\mathcal{A}$  的 t-SVD。

**证明** 由引理 1 以及  $\overline{\mathbf{B}}, \overline{\mathbf{C}}$  的块对角结构可知, 问题(2)可以分解为以下子问题

$$\min_{\overline{\mathcal{V}}_{(i, i, k)} \overline{\mathcal{V}}_{(i, i, k)}^T = \mathbf{I}} \langle \overline{\mathcal{A}}_{(i, i, k)}, \overline{\mathcal{V}}_{(i, i, k)} \rangle, k = 1, 2, \dots, s \quad (3)$$

由引理 2 知, 式(3)中各子问题的全局解存在,  $\overline{\mathcal{V}}_{(:, :, k)} = \overline{\mathcal{B}}_{(:, :, k)} \overline{\mathcal{D}}_{(:, :, k)}^T$ ,  $\overline{\mathcal{A}}_{(:, :, k)} = \overline{\mathcal{B}}_{(:, :, k)} \overline{\mathcal{C}}_{(:, :, k)} \cdot \overline{\mathcal{D}}_{(:, :, k)}^T$  是  $\overline{\mathcal{A}}_{(:, :, k)}$  的 t-SVD。由引理 1 中的  $\mathcal{F} = \mathcal{A} * \mathcal{B}$ , 则  $\overline{\mathbf{F}} = \overline{\mathbf{A}} \overline{\mathbf{B}}$  可知,  $\overline{\mathcal{V}} = \overline{\mathcal{B}} * \overline{\mathcal{D}}^T$ ,  $\mathcal{V} = \mathcal{B} * \mathcal{D}^T$ 。因此  $\mathcal{A} = \mathcal{B} * \mathcal{C} * \mathcal{D}^T$  是  $\mathcal{A}$  的 t-SVD。

## 2 E-TRPCA 方法

### 2.1 基于 t-SVD 的 E-TRPCA 模型

如图 1 所示, 矩阵可以分解为两个矩阵的积<sup>[12]</sup>

$$\mathbf{X} = \mathbf{U} \mathbf{V}^T + \mathbf{S} \quad (4)$$

借助 t 积, 可以把矩阵的情形推广到张量情形

$$\mathcal{X} = \mathcal{L} + \mathcal{e} = \mathcal{U} * \mathcal{V}^T + \mathcal{e} \quad (5)$$

事实上,如果对  $\mathcal{V}$  施加一个正交约束,则有  $\mathcal{U} = \mathcal{L} * \mathcal{V}$  与传统的核范数最小化 (nuclear norm minimization, NNM) 问题相同,对分量  $\mathcal{U}$  施加低秩约束,则可以构造如下 E-TNN 正则项

$$\begin{aligned} \|\mathcal{L}\|_{\text{E-TNN}} &= \min_{\mathcal{V} \in \mathbb{R}^{w \times r \times s}} \|\mathcal{L} * \mathcal{V}\|_* \\ \text{s. t. } \|\mathcal{L} * \mathcal{V}\|_F &= \|\mathcal{L}\|_F \\ \mathcal{V}^T * \mathcal{V} &= \mathcal{I} \end{aligned} \quad (6)$$

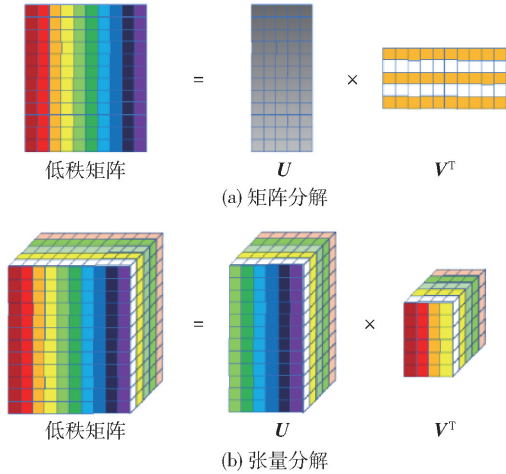


图1 低秩分解

Fig. 1 Low-rank factorization

第一个是约束在变换结果  $\mathcal{L} * \mathcal{V}$  上的 Frobenius 范数,这有助于避免变换引起的信息损失;第二个是对  $\mathcal{V}$  的正交约束,倾向于使变换后的结果  $\mathcal{L} * \mathcal{V}$  尽可能保留低秩张量  $\mathcal{L}$  的信息,此外它还可以帮助获得该变量的闭式解。由于不容易直接求解上述问题,因此将式(6)重新表述为以下等价问题

$$\begin{aligned} \|\mathcal{L}\|_{\text{E-TNN}} &= \min_{\mathcal{U}, \mathcal{V}} \|\mathcal{U}\|_* \\ \text{s. t. } \mathcal{L} &= \mathcal{U} * \mathcal{V}^T, \mathcal{V}^T * \mathcal{V} = \mathcal{I} \\ \mathcal{U} &\in \mathbb{R}^{h \times r \times s}, \mathcal{V} \in \mathbb{R}^{w \times r \times s} \end{aligned} \quad (7)$$

与 TNN 一样, E-TNN 同样作用在分量  $\mathcal{L}$  上,不同之处在于 E-TNN 不直接约束  $\mathcal{L}$  本身,而是约束从  $\mathcal{L}$  中学习得到的一组基,这组基是  $\mathcal{L}$  的所有侧切面的一个线性表示<sup>[9]</sup>。矩阵分解可以把矩阵分解为字典和稀疏矩阵的积,因此,借助上述张量工具,把分量  $\mathcal{L}$  分解为字典张量  $\mathcal{U}$  和投影张量  $\mathcal{V}$  的共轭转置的乘积。字典张量  $\mathcal{U}$  由分量  $\mathcal{L}$  经过正交变换得到:  $\mathcal{U} = \mathcal{L} * \mathcal{V}$ ,  $\mathcal{V}$  是正交变换张量,维度为  $w \times r \times s$ ,  $r < w$ 。另外,字典张量  $\mathcal{U}$  的规模总是小于低秩分量  $\mathcal{L}$ ,这使得它们受到噪声损坏的影响较小,从而可以获得更鲁棒的恢复效果。

将 E-TNN 正则项嵌入到 TRPCA 模型,得到如

下低秩张量恢复模型

$$\begin{aligned} \min_{\mathcal{U}, \mathcal{V}, \mathcal{L}, \mathcal{e}} \quad & \|\mathcal{U}\|_* + \lambda \|\mathcal{e}\|_1 \\ \text{s. t. } \quad & \mathcal{X} = \mathcal{L} + \mathcal{e} \\ & \mathcal{L} = \mathcal{U} * \mathcal{V}^T, \mathcal{V}^T * \mathcal{V} = \mathcal{I} \\ & \mathcal{U} \in \mathbb{R}^{h \times r \times s}, \mathcal{V} \in \mathbb{R}^{w \times r \times s} \end{aligned} \quad (8)$$

式中  $\mathcal{X}, \mathcal{L}, \mathcal{e} \in \mathbb{R}^{h \times w \times s}$  分别为观测张量、低秩张量和稀疏张量,  $l_1$  范数来约束稀疏张量,  $\lambda$  为超参数。

## 2.2 ADMM 优化

本节中,通过 ADMM 算法<sup>[13]</sup>求解模型(8)。由模型(8)得到以下增广拉格朗日函数。

$$\begin{aligned} \mathcal{L}(\mathcal{U}, \mathcal{V}, \mathcal{L}, \mathcal{e}, y_1, y_2) &= \|\mathcal{U}\|_* + \lambda \|\mathcal{e}\|_1 + \\ & \frac{\mu}{2} \left\| \mathcal{X} - \mathcal{L} - \mathcal{e} + \frac{y_1}{\mu} \right\|_F^2 + \frac{\mu}{2} \left\| \mathcal{L} - \mathcal{U} * \mathcal{V}^T + \frac{y_2}{\mu} \right\|_F^2 \end{aligned} \quad (9)$$

式中,  $\mu$  为惩罚项系数。ADMM 可以把上述问题分解为如下的 5 个子问题,在每一个子问题中,固定其余变量不变,只更新一个变量。

**$\mathcal{L}$  子问题** 从式(9)中找出所有包含  $\mathcal{L}$  的项,得到如下更新公式

$$\mathcal{L} = -\frac{1}{2} \left( \mathcal{X} - \mathcal{e} + \frac{y_1}{\mu} + \mathcal{U} * \mathcal{V}^T - \frac{y_2}{\mu} \right) \quad (10)$$

**$\mathcal{U}$  子问题** 从式(9)中找出所有包含  $\mathcal{U}$  的项,得到如下问题

$$\begin{aligned} \mathcal{U}^* &= \min_{\mathcal{U}} \|\mathcal{U}\|_* + \frac{\mu}{2} \left\| \mathcal{L} - \mathcal{U} * \mathcal{V}^T + \frac{y_2}{\mu} \right\|_F^2 = \\ \min_{\mathcal{U}} \quad & \|\mathcal{U}\|_* + \frac{\mu}{2} \left\| \mathcal{U} - \left( \mathcal{L} + \frac{y_2}{\mu} \right) \mathcal{V} \right\|_F^2 \end{aligned} \quad (11)$$

这个子问题可以通过张量奇异值阈值算子<sup>[2]</sup>求解

$$\mathcal{U}_* = \mathcal{D}_{\frac{1}{\mu}} \left( \left( \mathcal{L} + \frac{y_2}{\mu} \right) \mathcal{V} \right) \quad (12)$$

**$\mathcal{V}$  子问题** 从式(9)中找出所有包含  $\mathcal{V}$  的项,令  $\mathcal{P} = \mathcal{L} + \frac{y_2}{\mu}$ , 得到如下问题

$$\begin{aligned} \mathcal{V}^* &= \min_{\mathcal{V}} \frac{\mu}{2} \left\| \mathcal{L} - \mathcal{U} * \mathcal{V}^T + \frac{y_2}{\mu} \right\|_F^2 = \min_{\mathcal{V}} \frac{\mu}{2} \langle \mathcal{U} * \\ & \mathcal{V}^T - \mathcal{P} ( \mathcal{U} * \mathcal{V}^T - \mathcal{P} )^T \rangle = \min_{\mathcal{V}} \text{tr} ( \mathcal{P}^T * \mathcal{U} * \\ & \mathcal{V}^T )_{(:, :, 1)} = \min_{\mathcal{V}} \langle \mathcal{P}^T * \mathcal{U}, \mathcal{V} \rangle \end{aligned} \quad (13)$$

由定理 1 可知

$$\begin{cases} [\mathcal{B}, \sim, \mathcal{C}] = \text{t-svd}(\mathcal{P}^T * \mathcal{U}) \\ \mathcal{V}^* = \mathcal{B} * \mathcal{C}^T \end{cases} \quad (14)$$

**$\mathcal{e}$  子问题** 从式(9)中找出所有包含  $\mathcal{e}$  的项,得到如下问题

$$e^* = \min_e \left\| e \right\|_1 + \frac{\mu}{2} \left\| \mathcal{R} - \mathcal{L} - e + \frac{\mathcal{U}_1}{\mu} \right\|_F^2 \quad (15)$$

这个子问题可以通过软阈值收缩算子<sup>[14]</sup>来得到闭式解

$$e^* = e_{\frac{1}{\mu}} \left( \mathcal{R} - \mathcal{L} + \frac{\mathcal{U}_1}{\mu} \right) \quad (16)$$

拉格朗日乘子  $\mathcal{Y}_1, \mathcal{Y}_2$  通过如下公式来更新

$$\begin{cases} \mathcal{Y}_1 = \mathcal{Y}_1 + \mu(\mathcal{R} - \mathcal{L} - e) \\ \mathcal{Y}_2 = \mathcal{Y}_2 + \mu(\mathcal{L} - \mathcal{U} \mathcal{V}^T) \end{cases} \quad (17)$$

综上所述,求解 E-TRPCA 模型(8)的 ADMM 算法(算法1)详细过程如下。

输入:  $\mathcal{R} \in \mathbb{R}^{h \times w \times s}, r, \lambda, \mu = 10^{-2}, \mu_{\max} = 10^6, p = 1.3, e = 10^7$

输出:  $\mathcal{L}, e, \mathcal{U}, \mathcal{V}$

初始化参数: 初始化  $\mathcal{L}, e, \mathcal{Y}_1, \mathcal{Y}_2 = 0 \in \mathbb{R}^{h \times w \times s}, \mathcal{U} \in \mathbb{R}^{h \times w \times s}, \mathcal{V} \in \mathbb{R}^{w \times r \times s}$

1: while 不满足收敛条件

2: 通过式(10)更新  $\mathcal{L}$

3: 通过式(12)更新  $\mathcal{U}$

4: 通过式(14)更新  $\mathcal{V}$

5: 通过式(16)更新  $e$

6: 检查收敛条件

$$\frac{\left\| \mathcal{R} - \mathcal{L} - e \right\|_F^2}{\left\| \mathcal{R} \right\|_F^2} \leq e, \left\| \mathcal{L} - \mathcal{U} * \mathcal{V}^T \right\|_F^2 \leq e$$

7: 通过式(17)更新  $\mathcal{Y}_1, \mathcal{Y}_2$

8: end while

### 2.3 复杂度分析

算法1以三阶张量  $\mathcal{R} \in \mathbb{R}^{h \times w \times s}$  作为输入,主要的计算成本是在更新  $\mathcal{U}, \mathcal{V}$  时产生的。更新  $\mathcal{U}$  需要计算一个大小为  $h \times w \times s$  张量的 t-SVD,更新  $\mathcal{V}$  时需要计算一个大小为  $w \times r \times s$  张量的 t-SVD,那么算法1的时间复杂度为  $O(\max(h, w)rs \log s + (\max(h, w)r^2s))$ 。

## 3 实验验证

以图像恢复作为仿真实验、背景建模作为真实实验来测试所提方法,并与 RPCA 方法<sup>[1]</sup>、基于 Tucker 分解的 SNN 方法<sup>[4]</sup>、基于 t-SVD 的 TNN 方法<sup>[2]</sup>、基于张量火车分解的 TTNN 方法<sup>[6]</sup>和基于张量环分解的 RTRC 方法<sup>[7]</sup>这5种方法进行对比。其中,RTRC 方法被用来解决张量补全问题,在本文中用于张量鲁棒主成分分析,记为 TRNN(tensor ring nuclear norm)。图像恢复实验中的图片和背景建模

中的视频分别用三阶张量和四阶张量存储,它们的值均被缩放到  $[0, 1]$ 。

### 3.1 图像恢复

使用伯克利分割数据集<sup>[15]</sup>中的彩色图像进行测试。图片的大小为  $321 \times 481 \times 3$  或  $481 \times 321 \times 3$ 。RPCA、TNN、TRNN 和本文方法可以直接处理三阶张量,TTNN 方法在数据输入模型之前采用 ketaugmentation (KA)技术作了数据增强处理,因此需要把图片大小调整为  $320 \times 480 \times 3$ ,再利用 KA 方法将图片尺寸转化为  $[4 \times 4 \times 4 \times 5 \times 4 \times 4 \times 5 \times 6 \times 3]$ 。

**噪声模拟** 本文主要去除图像中的随机噪声,随机噪声比例  $p$  分别设为 0.1、0.2、0.3,以  $p=0.1$  为例,即图像有 10% 的像素被设为区间  $[0, 1]$  之间的一个随机值,被破坏的像素的位置是随机的。

**评价指标** 为了评价比较方法去除随机噪声的性能,选取峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似度 (structural similarity index, SSIM) 作为评价指标<sup>[16]</sup>,PSNR 值和 SSIM 值越高,表示恢复结果越好。

**参数分析** 本文提出的模型包含正则项参数  $\lambda$  和 tubal 秩  $r$ 。正则项参数  $\lambda$  用来控制随机噪声对恢复结果的影响。由图2知,当  $p=0.1$  时,在  $\lambda$  为 0.045 处 PSNR 取得最大值,  $p=0.2$  时,在  $\lambda$  为 0.041 处取得最大值,  $p=0.3$  时,在  $\lambda$  为 0.032 处取得最大值。  $p=0.1$  时,本文设定  $\lambda = \frac{1}{\sqrt{\min(h, w)s}}$ ,

$p=0.2, 0.3$  时,本文设定  $\lambda = \frac{1}{\sqrt{\min(h, w)s}}$ 。

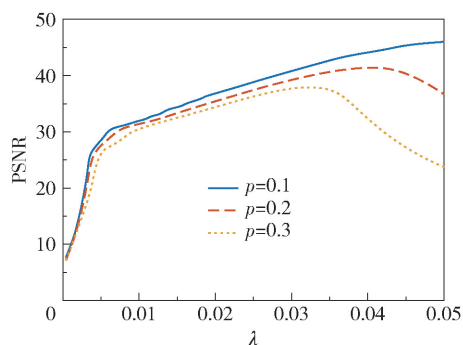


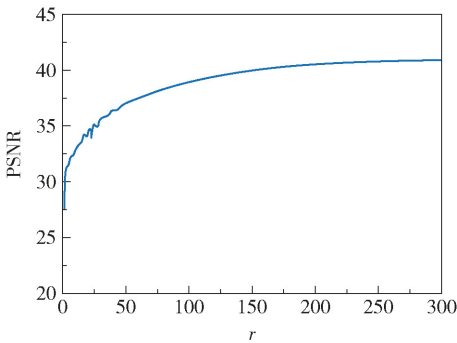
图2  $p=0.1, 0.2, 0.3$  时, PSNR 值随  $\lambda$  的变化曲线  
Fig.2 PSNR values changing with  $\lambda$  for  $p=0.1, 0.2, 0.3$

Tubal 秩  $r$  用于刻画图像的低秩性。本文通过遍历  $r$  来分析 tubal 秩对 PSNR 的影响。由图3(b)可知,随着  $r$  的增大,图像的恢复效果也越来越好。





(a) 示例图片5



(b) PSNR曲线

图 3 管道秩对 PSNR 值的影响

Fig. 3 The effect of tubal rank on PSNR value

为了在效率和性能之间取得一个平衡,将  $r$  设为  $\frac{\min(h,w)}{2}$ 。

1) 视觉效果比较

为了直观地对比图像的恢复结果,从伯克利分割数据集中选取了 6 张示例图片进行可视化展示。从图 4 ~ 9 中可以看出,RPCA、SNN、TTNN 和 TRNN

算法虽然能够去除图片中的随机噪声和椒盐噪声,但这 4 种方法不能很好地保留图像中的细节。图 4(h) ~ 9(h) 展示了本文所提模型的恢复结果,可以看到添加的噪声被去除,并且图像中的纹理和边缘也得到了很好的保留。这说明本文提出的 E-TNN 正则项要比 t-SVD 张量框架下的张量核范数更有效。

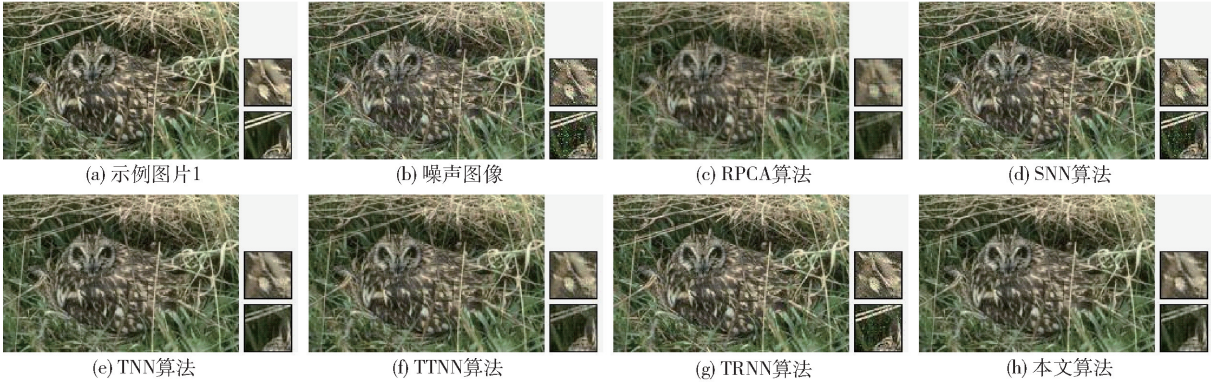


图 4 所有对比方法在示例图片 1 下的恢复结果( $p=0.1$ )

Fig. 4 Restoration results of all competing methods for example image 1 ( $p=0.1$ )

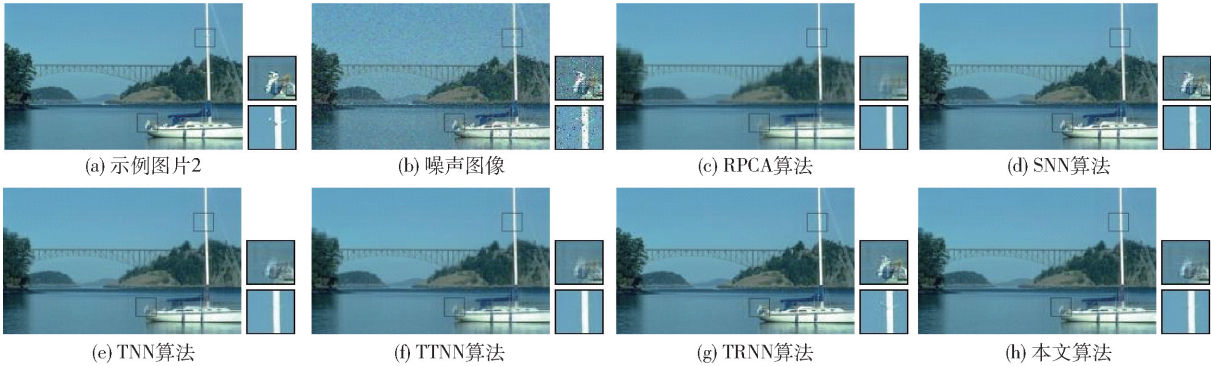


图 5 所有对比方法在示例图片 2 下的恢复结果( $p=0.1$ )

Fig. 5 Restoration results of all competing methods for example image 2 ( $p=0.1$ )



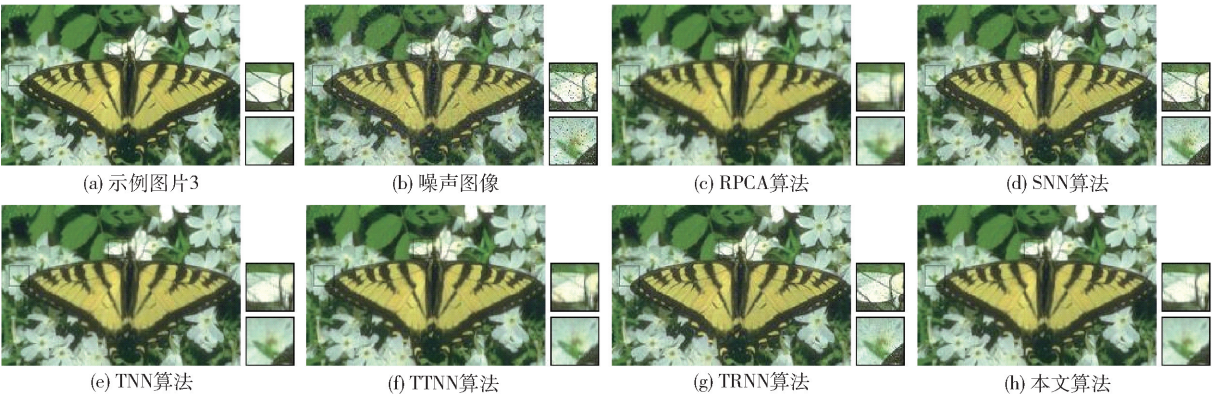


图 6 所有对比方法在示例图片 3 下的恢复结果( $p=0.1$ )

Fig. 6 Restoration results of all competing methods for example image 3( $p=0.1$ )

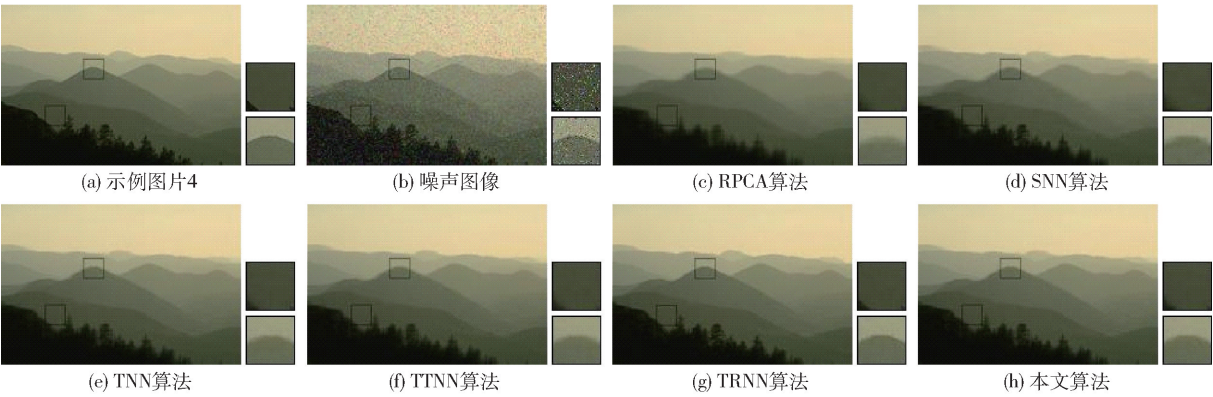


图 7 所有对比方法在示例图片 4 下的恢复结果( $p=0.1$ )

Fig. 7 Restoration results of all competing methods for example image 4( $p=0.1$ )

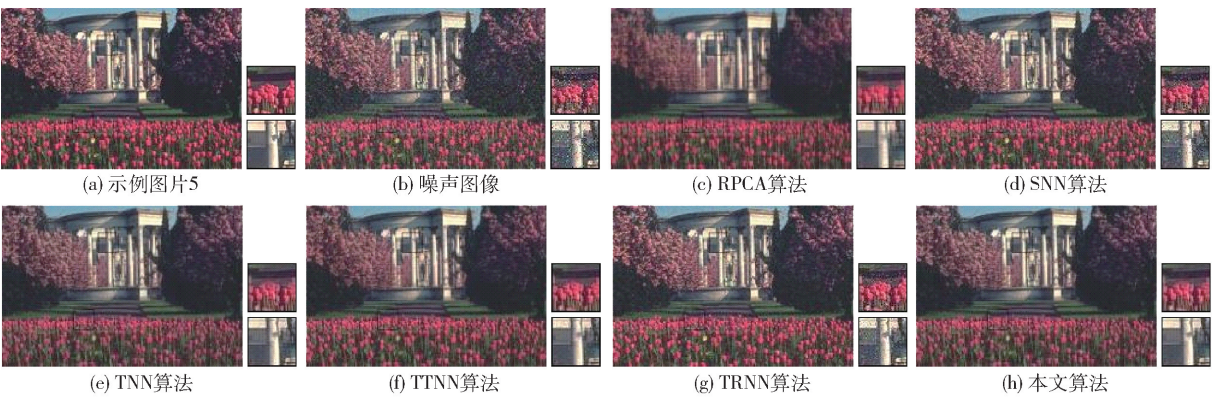


图 8 所有对比方法在示例图片 5 下的恢复结果( $p=0.1$ )

Fig. 8 Restoration results of all competing methods for example image 5( $p=0.1$ )

2) 定量结果比较

表 1~3 分别是 6 张示例图片在不同随机噪声比例  $p=0.1, 0.2, 0.3$  下的 PSNR 值和 SSIM 值。可以看出,本文所提模型的 PSNR 评价指标在所有的噪声比例中都取得了最好的结果,而且在 SSIM 评价指标上也基本上取得了最优或次优的结果。

图 10 给出了在  $p=0.1$  时伯克利分割数据集 50

张图片的 PSNR 值和运行时间的比较。可以看出,与其他几种算法相比,本文所提方法取得的 PSNR 值最高,并且运行所需时间也仅次于 TRNN 方法,远低于其他方法。

此外,由表 1~3 可知,在示例图片 1 和示例图片 5 中,所有方法都没有取得比较好的恢复效果,而示例图片 4 和示例图片 6 的恢复效果较好。

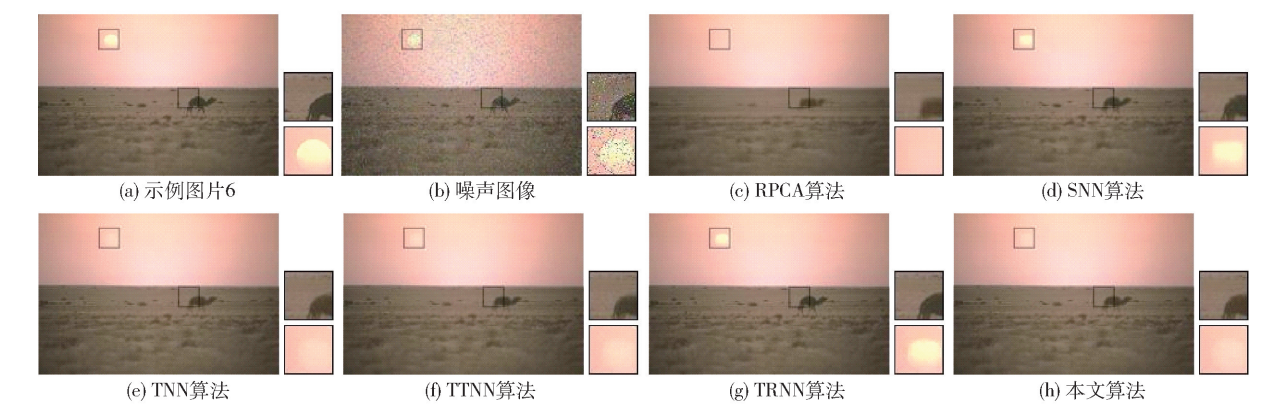


图 9 所有对比方法在示例图片 6 下的恢复结果( $p=0.1$ )

Fig. 9 Restoration results of all competing methods for example image 6( $p=0.1$ )

表 1 各对比方法在  $p=0.1$  下的 PSNR 和 SSIM 值比较

Table 1 Comparison of the PSNR and SSIM values obtained using all competing methods with  $p=0.1$

方法	示例图片 1		示例图片 2		示例图片 3		示例图片 4		示例图片 5		示例图片 6	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RPCA	19.782	0.754 3	27.065	0.882 9	24.226	0.815 4	35.297	0.930 8	21.769	0.749 9	37.30	0.952
SNN	20.343	0.732 9	30.526	0.874 4	22.313	0.548 9	35.909	0.914 8	21.766	0.648 6	38.62	0.940 3
TNN	<b>25.607</b>	<i>0.941 7</i>	31.032	<b>0.954 2</b>	<b>28.040</b>	<i>0.924 3</i>	<b>38.024</b>	0.966 7	<b>24.891</b>	<i>0.893 7</i>	41.40	0.984 7
TTNN	21.984	0.787 9	<b>35.292</b>	0.949 5	24.544	0.633 2	41.186	<i>0.978 0</i>	23.110	0.696 6	<b>45.16</b>	<i>0.987 7</i>
TRNN	23.399	0.890 3	30.022	0.936 2	27.226	0.892 6	37.431	0.956 6	24.554	0.864 8	40.26	0.976 1
E-TNN	26.186	<b>0.911 9</b>	<i>36.112</i>	<i>0.962 4</i>	<i>30.934</i>	<b>0.909 4</b>	<i>41.717</i>	<b>0.971 5</b>	27.787	<b>0.883 8</b>	<i>45.961</i>	<b>0.986 8</b>

斜体表示最好结果,黑色加粗表示次优结果。

表 2 各对比方法在  $p=0.2$  下的 PSNR 和 SSIM 值比较

Table 2 Comparison of the PSNR and SSIM values obtained using all competing methods with  $p=0.2$

方法	示例图片 1		示例图片 2		示例图片 3		示例图片 4		示例图片 5		示例图片 6	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RPCA	18.615	0.657 1	26.095	0.853 2	22.675	0.715 9	34.321	0.916 5	20.950	0.683 0	35.912	0.938 1
SNN	15.693	0.506 9	17.235	0.197 0	15.751	0.274 0	16.844	0.089 9	15.958	0.385 9	17.445	0.106 5
TNN	<b>22.974</b>	<b>0.876 6</b>	<b>29.320</b>	<b>0.926 6</b>	<b>25.948</b>	<b>0.849 9</b>	<b>36.637</b>	<b>0.951 9</b>	<b>23.260</b>	<i>0.823 8</i>	<b>39.295</b>	<b>0.972 8</b>
TTNN	16.658	0.548 4	21.036	0.312 6	17.514	0.326 7	30.925	0.656 2	17.517	0.442 9	34.955	0.836 4
TRNN	20.639	0.761 6	28.167	0.867 8	24.449	0.722 9	35.940	0.925 8	22.637	0.744 9	37.588	0.948 5
E-TNN	23.060	<i>0.877 6</i>	<i>30.481</i>	<i>0.933</i>	<i>26.721</i>	<i>0.853 6</i>	37.258	<i>0.952 0</i>	23.865	<b>0.820 2</b>	<i>40.172</i>	<i>0.974 2</i>

斜体表示最好结果,黑色加粗表示次优结果。

表 3 各对比方法在  $p=0.3$  下的 PSNR 和 SSIM 值比较

Table 3 Comparison of the PSNR and SSIM values obtained using all competing methods with  $p=0.3$

方法	示例图片 1		示例图片 2		示例图片 3		示例图片 4		示例图片 5		示例图片 6	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RPCA	17.390	0.541 1	24.953	0.799 8	20.624	0.882 6	32.936	0.916 5	19.897	0.882 6	34.286	0.912 3
SNN	13.592	0.381 5	14.340	0.119 2	13.365	0.043 2	13.529	0.089 9	13.578	0.043 2	14.065	0.053
TNN	<b>20.546</b>	<i>0.761 9</i>	<b>27.590</b>	<i>0.878 8</i>	<b>23.294</b>	<i>0.923 1</i>	<b>34.908</b>	<i>0.951 9</i>	<b>21.537</b>	<i>0.923 1</i>	<b>37.282</b>	<i>0.952 1</i>
TTNN	14.079	0.400 5	16.167	0.152 5	14.209	0.074 3	16.734	0.656 2	14.589	0.074 3	17.667	0.095 6
TRNN	17.892	0.591 5	25.332	0.700 1	20.249	0.792	32.380	<b>0.925 8</b>	19.993	0.792	33.925	0.859 5
E-TNN	20.650	<b>0.749 9</b>	28.449	<b>0.868 9</b>	23.616	<b>0.913 1</b>	35.265	0.904 6	21.777	<b>0.913 1</b>	<i>37.801</i>	<b>0.948 7</b>

斜体表示最好结果,黑色加粗表示次优结果。



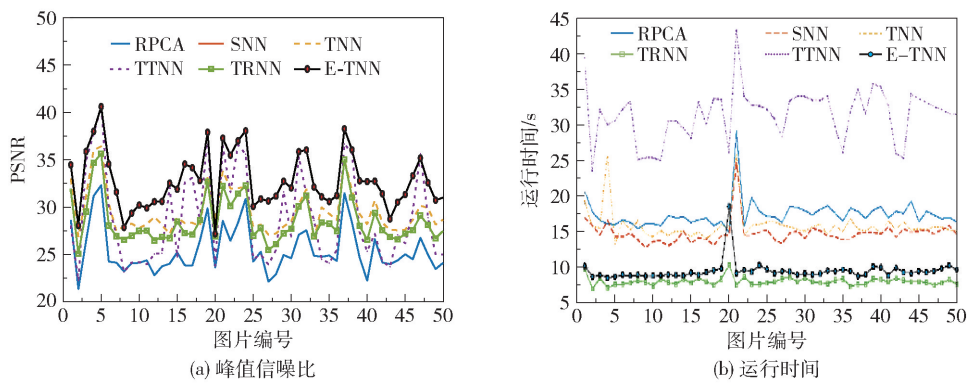
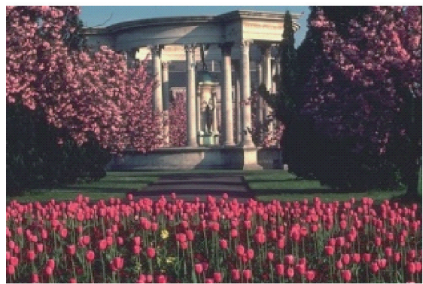


图 10 各对比方法在伯克利分割数据集 50 张图片上的定量结果比较

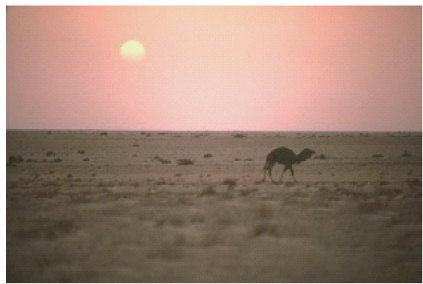
Fig. 10 Quantitative comparison of all competing methods on 50 images of the Berkeley segmentation dataset

RPCA、SNN、TNN、TTNN、TRNN 等方法都假设数据具有低秩结构,对低秩性有强依赖性。我们发现,对于结构比较复杂、纹理较为丰富的示例图片 5(图 11(a)),其张量奇异值并没有迅速接近于 0(图 11(b)),这表明其 tubal 秩相对较大,上述方法难以取得好的恢复效果。而在背景空旷的示例图片 6(图 11(c))中,其奇异值则迅速接近于 0(图 11(d)),表明 tubal 秩较小,上述方法可以取得较好的恢复效果。对此,本文新增了一个参数  $r$  用于控制 tubal 秩对数据恢复效果的影响,E-TRPCA 模型中的  $\lambda$  参数可以协调核范数正则项和  $l_1$  正则项,参数  $r$  则可以进一步控制核范数正则

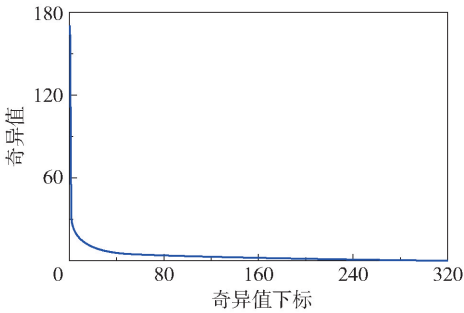
项。由图 3 可知,当随机噪声比例  $p=0.1$  时,在不改变其他参数的情况下, $r$  从 0 增加 300,PSNR 也从 25 增加到 42,这为处理不同场景的图像提供了一个优化手段。为了更好地评估本文模型的泛化性,图 12 给出了  $r=150$  和  $r=300$  两种情况下本文方法与其他 5 种方法的 PSNR 值比较。从图中可以看出,当  $r$  取 150 和 300 时,本文所提方法取得的 PSNR 值都要高于 RPCA、SNN、TNN、TTNN 和 TRNN 方法,而且  $r=300$  时所得的结果要好于  $r=150$  时的结果。这说明在数据是否具有低秩性的先验信息未知的情况下,本文所提方法具有较好的泛化性。



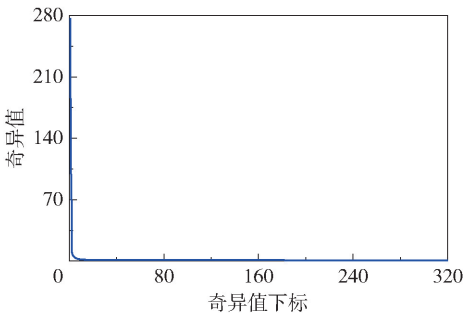
(a) 示例图片 5



(c) 示例图片 6



(b) 图(a)的奇异值



(d) 图(c)的奇异值

图 11 不同图片的奇异值分布

Fig. 11 Singular value distribution of different images



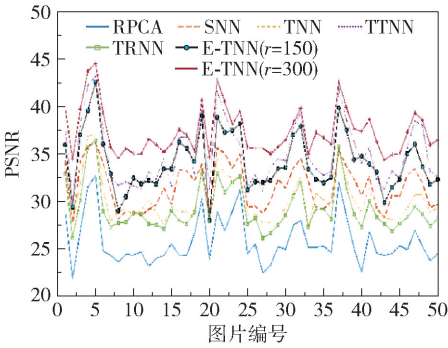


图 12  $r = 150, 300$  时, 50 张图片的 PSNR 值  
Fig. 12 PSNR values of 50 images for  $r = 150, 300$

3.2 背景建模

视频的前景和背景分离任务一直以来都是研究的热点。视频由连续的帧图像序列组成, 其中结构稳定、变动很小的场景内容是视频的背景。由于背景之间的相关性, 可以将其视作一个低秩分量, 而在视频中所占像素较小且变动显著的物体, 如汽车或行人, 即为视频的前景, 可以视为稀疏分量。本文选取了 driveway(117 帧)<sup>[17]</sup> 和 shop(100 帧)<sup>[18]</sup> 两个视频。在 RPCA 算法中, 将视频张量  $h \times w \times 3 \times f$  展开为矩阵  $hw \times 3f$ ; 在 SNN、TNN、TRNN 算法中, 将视频张量的大小转化为  $hw \times 3 \times f$ ; 在本文所提方法中, 将视频张量转化为  $hw \times f \times 3$ ,  $f$  表示视频的帧数。

表 4 为 RPCA、SNN、TNN、TRNN 和本文方法运行时间的比较, 图 13~16 给出了 5 种方法在两个视频上的结果 (TTNN 方法难以找到一个合适的 KA 参数, 故不参与比较)。可知 5 种方法都可以将运动的行人从背景中提取出来, 但本文方法的用时

表 4 运行时间比较

视频	运行时间/s				
	RPCA	SNN	TNN	TRNN	E-TNN
driveway	500.8	809.9	401.3	379.2	285.4
shop	119.2	386.5	98.8	90.3	86.7

最短。  
3.3 去除高斯噪声

本文在原有模型的基础上新增了一个正则项  $\|\mathcal{N}\|_{21}$  来刻画高斯噪声。通过最小化核范数、 $l_1$  范数和  $l_{21}$  的组合, 分离出被混合噪声破坏的干净数据。

$l_{21}$  范数定义如下。

$$\|\mathbf{A}\|_{21} = \sum_k \sum_j \sqrt{\sum_i |a_{ijk}|^2}, \text{ 这里 } \mathbf{A} = (a_{ijk})_{n_1 \times n_2 \times n_3}.$$

此时, 模型(8)可以改写为

$$\begin{aligned} \min_{\mathcal{U}, \mathcal{V}, \mathcal{L}, e} \quad & \|\mathcal{U}\|_* + \lambda \|e\|_1 + \beta \|\mathcal{N}\|_{21} \\ \text{s. t.} \quad & \mathcal{X} = \mathcal{L} + e + \mathcal{N} \\ & \mathcal{L} = \mathcal{U} * \mathcal{V}^T, \mathcal{V}^T * \mathcal{V} = \mathcal{I} \\ & \mathcal{U} \in \mathbb{R}^{h \times r \times s}, \mathcal{V} \in \mathbb{R}^{w \times r \times s} \end{aligned} \tag{18}$$

模型(18)的增广拉格朗日函数为

$$\begin{aligned} \mathcal{L}(\mathcal{U}, \mathcal{V}, \mathcal{L}, e, y_1, y_2) = & \|\mathcal{U}\|_* + \lambda \|e\|_1 + \beta \|\mathcal{N}\|_{21} + \frac{\mu}{2} \left\| \mathcal{X} - \mathcal{L} - e - \mathcal{N} + \frac{y_1}{\mu} \right\|_F^2 + \frac{\mu}{2} \left\| \mathcal{L} - \mathcal{U} * \mathcal{V}^T + \frac{y_2}{\mu} \right\|_F^2 \end{aligned} \tag{19}$$

式(19)依然采取 ADMM 算法求解, 与算法 1 相

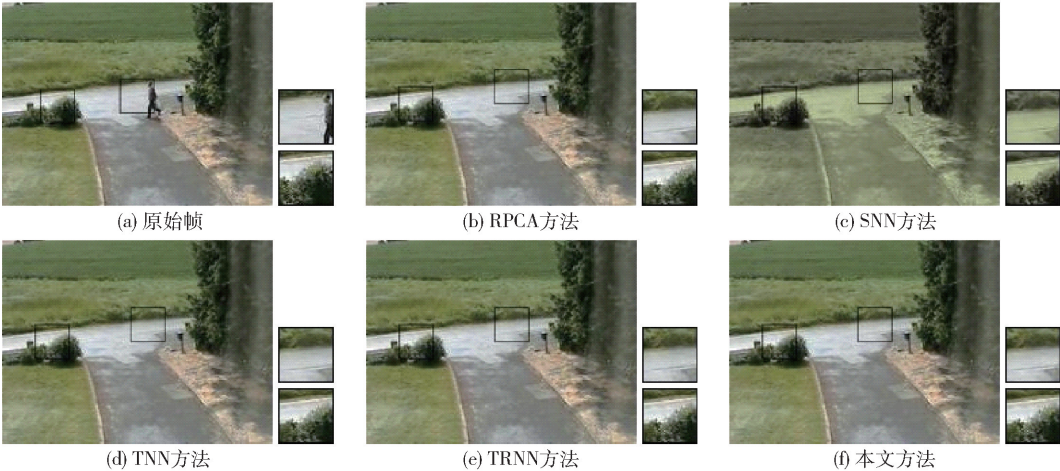


图 13 driveway 视频的背景图像  
Fig. 13 Background image for driveway video

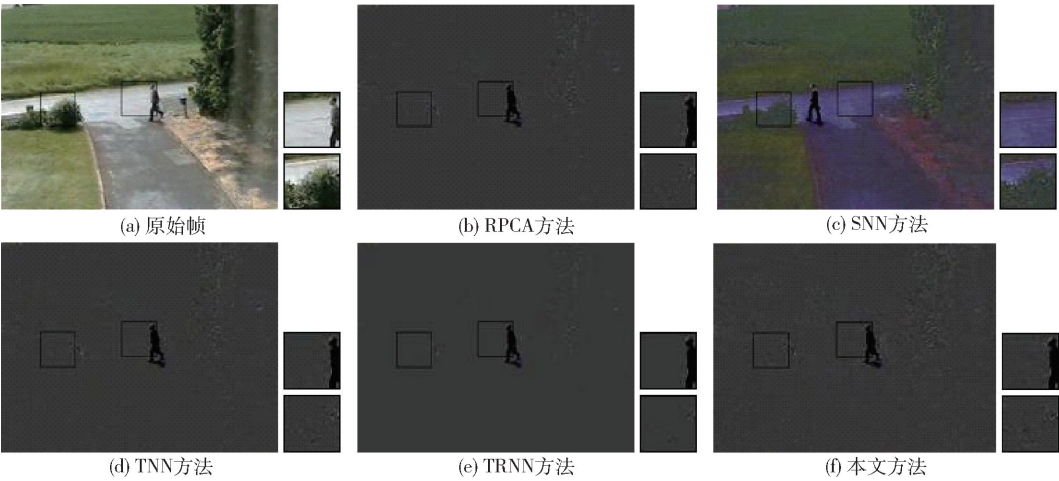


图 14 driveway 视频的前景图像  
Fig. 14 Foreground image for driveway video

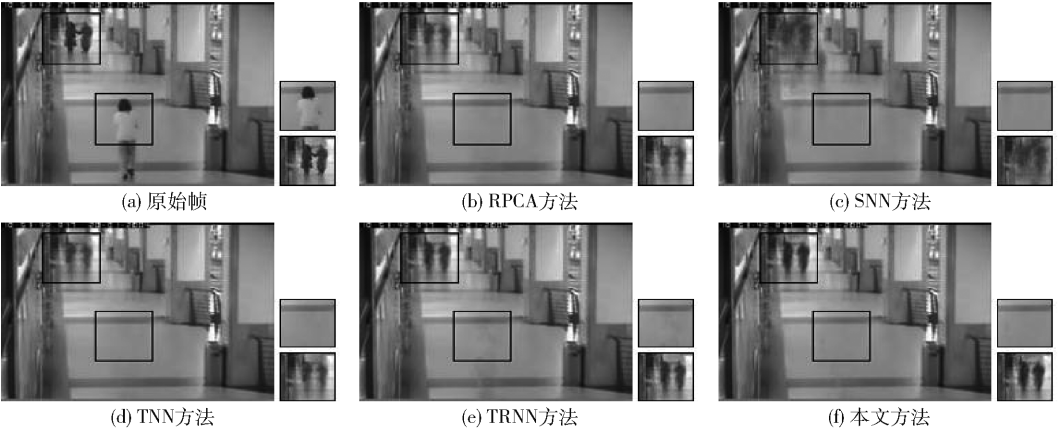


图 15 shop 视频的背景图像  
Fig. 15 Background image for shop video

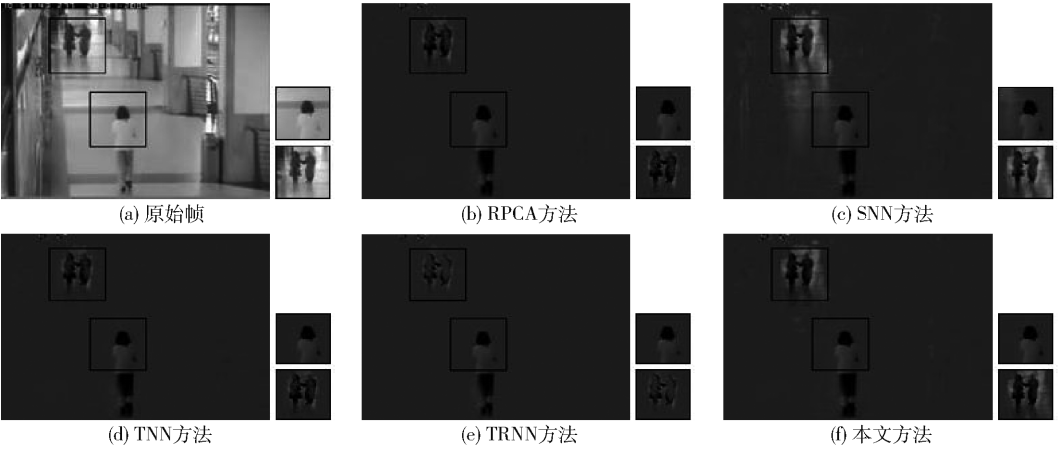


图 16 shop 视频的前景图像  
Fig. 16 Foreground image for shop video

比,求解模型(8)多出一个子问题,即求解  $\|\mathcal{N}\|_{21}$  子问题,该子问题可以写成如下格式。

$$\mathcal{N}^* = \min_{\mathcal{N}} \|\mathcal{N}\|_{21} + \frac{\mu}{2} \left\| \mathcal{N} - \left( \mathcal{X} - \mathcal{L} - \mathcal{e} + \frac{\mathcal{L}_1}{\mu} \right) \right\|_F^2 \quad (20)$$

令  $\mathcal{E} = \mathcal{X} - \mathcal{L} - \mathcal{e} + \frac{\mathcal{U}_1}{\mu}$ , 则有

$$\mathcal{N}_{(:,j,k)} = \begin{cases} \frac{\|\mathcal{E}_{(:,j,k)}\|_2 - \frac{\beta}{\mu}}{\|\mathcal{E}_{(:,j,k)}\|_2} \mathcal{E}_{(:,j,k)}, & \frac{\beta}{\mu} < \|\mathcal{E}_{(:,j,k)}\|_2 \\ 0, & \text{其他} \end{cases} \quad (21)$$

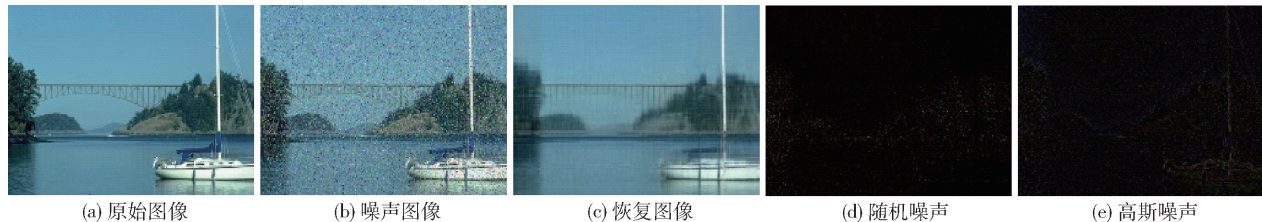


图 17 混合噪声去除

Fig. 17 Mixture noise removal

## 4 结束语

本文引入了增强的 TNN 正则项 (E-TNN), 提出一种增强的 TRPCA 模型。与 TNN 相比, E-TNN 能够在张量数据的子空间投影上计算低秩性, 可刻画张量数据中各成分的相关性和差异, 从而更真实地反映张量数据的潜在低秩结构。与张量鲁棒主成分分析方法相比, 本文所提算法可以有效地恢复出被噪声破坏的图像, 并缩短运行所需时间。下一步可以基于不同维度的张量数据自动调整字典张量的维度。此外, 基于广义线性变换的 t 积张量框架在张量低秩恢复问题中的应用将是未来的研究方向之一。

## 参考文献:

- [1] CANDÈS E J, LI X, MA Y, et al. Robust principal component analysis? [J]. Journal of the ACM, 2011, 58(3): 11.
- [2] LU C, FENG J, CHEN Y, et al. Tensor robust principal component analysis with a new tensor nuclear norm[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(4): 925–938.
- [3] TOCK D. Tensor decomposition and its applications[D]. Chester: University of Chester, 2010.
- [4] HUANG B, MU C, GOLDFARB D, et al. Provable models for robust low-rank tensor completion [J]. Pacific Journal of Optimization, 2015, 11(2): 339–364.
- [5] YANG J H, ZHAO X L, JI T Y, et al. Low-rank tensor train for tensor robust principal component analysis[J].

本文将模型(18)应用于图片恢复。图 17 给出了本文方法去除混合噪声的可视化结果。图 17(b)是添加了均值为 0、方差为 0.05 的高斯噪声和随机噪声比例  $p=0.2$  的示例图片 2。从图 17(c)~(e)可以看出, 图像的低秩部分、随机噪声部分和高斯噪声部分都被很好地分离出来。

Applied Mathematics and Computation, 2020, 367:124783.

- [6] DIAN R, LI S T, FANG L Y. Learning a low tensor-train rank representation for hyperspectral image super-resolution [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(9): 2672–2683.
- [7] HUANG H Y, LIU Y P, LONG Z, et al. Robust low-rank tensor ring completion [J]. IEEE Transactions on Computational Imaging, 2020, 6: 1117–1126.
- [8] KILMER M E, MARTIN C D. Factorization strategies for third-order tensors [J]. Linear Algebra and its Applications, 2011, 435(3): 641–658.
- [9] ZHOU P, LU C Y, FENG J S, et al. Tensor low-rank representation for data recovery and clustering [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(5): 1718–1732.
- [10] KILMER M E, BRAMAN K, HAO N, et al. Third-order tensors as operators on matrices: a theoretical and computational framework with applications in imaging [J]. SIAM Journal on Matrix Analysis and Applications, 2013, 34(1): 148–172.
- [11] XIE Q, ZHAO Q, MENG D Y, et al. Multispectral images denoising by intrinsic tensor sparsity regularization [C]//Proceedings of the IEEE Computer Society Conference on 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016: 1692–1700.
- [12] HE W, ZHANG H Y, ZHANG L P, et al. Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration [J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(1): 178–188.
- [13] LIN Z, CHEN M, MA Y. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank

- matrices[EB/OL]. (2013-10-18). arXiv: 1009.5055.
- [14] CAI J F, CANDÈS E J, SHEN Z. A singular value thresholding algorithm for matrix completion[J]. SIAM Journal on Optimization, 2010, 20(4): 1956-1982.
- [15] MARTIN D, FOWLKES C, TAL D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[C]//Proceedings of Eighth IEEE International Conference on Computer Vision. Vancouver, 2001: 416-423.
- [16] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [17] BARNICH O, VAN DROOGENBROECK M. ViBe: a universal background subtraction algorithm for video sequences[J]. IEEE Transactions on Image Processing, 2011, 20(6): 1709-1724.
- [18] BOUWMANS T, SOBRAL A, JAVED S, et al. Decomposition into low-rank plus additive matrices for background/foreground separation: a review for a comparative evaluation with a large-scale dataset[J]. Computer Science Review, 2017, 23: 1-71.

## Applications of enhanced tensor robust principal component analysis

ZHAO FengYing<sup>1</sup> YANG HongWei<sup>2</sup> ZHAO LiNa<sup>1\*</sup>

(1. College of Mathematics and Physics; 2. Center for Information Technology, Beijing University of Chemical Technology, Beijing 100029, China)

**Abstract:** Robust principal component analysis (RPCA) is a popular model to deal with image restoration and background modeling problems. By focusing on the problem of the excessive dependence of the original RPCA and an improved model based on the low-rank structure of the input data, we propose an enhanced robust principal component analysis model and design a new enhanced tensor nuclear norm. To naturally reflect the intrinsic structure of the tensor and improve the generalization of the model, E-TNN restricts the low-rank properties of the tensor data via its low dimensional subspace bases. The augmented Lagrange multiplier method is used to optimize the objective function. Experimental measurements of image denoising and background modeling show that the proposed method outperforms other current methods in terms of image restoration effect and running time.

**Key words:** tensor robust principal component analysis (TRPCA); tensor low-rank recovery; enhanced tensor nuclear norm; tensor decomposition

(责任编辑:吴万玲)