

推广的张量鲁棒主成分分析模型及其应用

王 颖, 唐科威

辽宁师范大学数学学院, 辽宁 大连

收稿日期: 2023年7月7日; 录用日期: 2023年8月8日; 发布日期: 2023年8月15日

摘 要

近年来, 张量在处理可视化数据方面有广泛的应用, 以鲁棒主成分分析(RPCA)为基础, 将其扩展至张量情况, 张量鲁棒主成分分析(TRPCA)被提出, 该模型已经成功应用于恢复彩色图像、视频的前背景分割等方面。然而TRPCA仅仅考虑了本身具有低秩性的图像, 不能校正倾斜的彩色图像, 为了解决这个问题, 本文通过考虑变换后张量的低秩性和稀疏性进行建模, 对TRPCA进行了推广, 同时, 我们还引入了张量的F范数来更好地处理高斯噪声和分割视频的动态背景。最后, 在不同类型的彩色图像和视频上进行了大量实验, 证明了本文方法的有效性。

关键词

张量, 核范数, 倾斜图像恢复, 高斯噪声, 动态背景分割

Extended Tensor Robust Principal Component Analysis Model and Its Application

Ying Wang, Kewei Tang

School of Mathematics, Liaoning Normal University, Dalian Liaoning

Received: Jul. 7th, 2023; accepted: Aug. 8th, 2023; published: Aug. 15th, 2023

Abstract

In recent years, tensors have been widely used in processing visualization data. To extend robust principal component analysis (RPCA) to tensor situations, tensor robust principal component analysis (TRPCA) has been proposed, and this model has been successfully applied to color images restoration, video background segmentation, and other aspects. However, TRPCA only considers images with low rank properties and cannot correct skewed color images. To solve this problem,

we extend TRPCA by considering the low rankness and sparsity of the transformed tensor. In addition, we also introduce F norm to better handle Gaussian noise and segment the dynamic background of video. Finally, a large number of experiments are conducted on different types of color images and videos to demonstrate the effectiveness of the proposed method.

Keywords

Tensor, Nuclear Norm, Skewed Image Recovery, Gaussian Noise, Dynamic Background Segmentation

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在计算机视觉和机器学习领域, 自从向量的稀疏性被提出[1] [2] [3]后, 该性质被广泛应用于大量模型, 至今仍被很多文章使用。向量稀疏性的获得需要最小化向量的 0 范数, 即向量非零分量的个数尽量少, 为了便于求解, 最常见的做法是最小化向量的 1 范数, 对于向量 $v \in \mathbb{C}^n$, 通常我们用 $\|v\|_1 = \sum_i |v_i|$ 来表示 1 范数。但是灰度图像本质是矩阵, 灰度视频本质是一个三阶张量, 这样的数据转化成向量后, 追求稀疏性不再具有明显的意义, 这就需要研究矩阵甚至张量的相关性质。

为了适用于更多的实际问题, 矩阵的低秩性被提出, 该性质也被广泛应用于大量的模型中, 追求矩阵低秩性的常见做法是最小化矩阵的核范数, 我们通常用 $\|M\|_* := \sum_i \sigma_i(M)$ 来表示矩阵的核范数, 即矩阵 M 的所有奇异值之和。经典的模型包括鲁棒的主成分分析(RPCA) [4] [5], 假设一个数据矩阵 X 是由一个低秩分量和稀疏分量叠加组成的, 即 $X = L_0 + E_0$, 其中 L_0 为具有低秩信息的低秩矩阵, E_0 为具有稀疏结构的稀疏矩阵。在文献[4]中表明, 如果满足某些条件, 则可以通过求解凸优化问题

$$\min_{L, E} \|L\|_* + \lambda \|E\|_1, \quad \text{s.t. } X = L + E,$$

来高概率地恢复 L_0 和 E_0 。其中 $\|E\|_1$ 表示矩阵 1 范数, 即 E 中所有元素的绝对值的和, $\lambda > 0$ 是一个加权参数, 用于权衡 1 范数和核范数两部分的影响。鲁棒主成分分析对噪声表现出高度的鲁棒性, 可以在噪声存在的情况精确稳定地实现低秩原始图像的恢复重构, 因此, 该模型及其扩展已经成功应用于人脸识别、图像的复原和视频的前背景分割[6]等场景。但是, 实际问题中, 图像在拍摄过程中的视角往往有一定的倾斜, 这样原始图像并不直接具备低秩性, 经过变换的图像才具备低秩性, 由此 Transform Invariant Low-rank Textures (TILT)模型[7]被提出, TILT 是通过仿射或射影变换, 利用变换后图像具备的低秩性进行建模, 从而可以处理图像摆正等问题。假设我们给出一个具有噪声的倾斜图像:

$$I = (I^0 + E) \circ \tau^{-1},$$

其中 I^0 是低秩的, E 是稀疏的, $\tau: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 属于某个变换群。在 TILT 中表明, 如果满足某些条件, 则可以通过求解凸优化问题

$$\begin{aligned} & \min_{I^0, E, \tau} \|I^0\|_* + \lambda \|E\|_1, \\ & \text{s.t. } I \circ \tau + \nabla I \Delta \tau = I^0 + E, \\ & A \Delta \tau = 0, \end{aligned}$$

高概率地恢复 I^0 和 E 。其中 ∇I 是关于变换 τ 参数的导数的雅克比矩阵, A 是 τ 的线性化约束。然而, RPCA 和 TILT 都只能处理矩阵数据, 真实图像的数据在本质上往往是多阶的, 例如, 一个 RGB 彩色图像是一个由红色、绿色和蓝色三个颜色通道组成的三阶张量; 一个灰度视频由两个空间变量和一个时间变量进行索引的三阶张量[8]。要使用 TILT, 首先必须将多阶数据转换为一个矩阵, 这种预处理通常会导致信息丢失, 并会导致性能下降。为了缓解这个问题, 很自然地考虑利用张量的某些性质进行建模。

近年来, 张量的低秩性被学者们广泛讨论, 卢等人提出了一种新的张量核范数并将 RPCA 模型进行了推广给出了张量的鲁棒主成分分析模型(TRPCA) [9], 由于张量低秩性的使用, 该模型在处理彩色图像和灰度视频时更加得心应手, 主要体现在 RPCA 在处理灰度视频时, 由于其只能考虑矩阵的性质, 所以需要将视频的每一帧都拉成向量再拼接成矩阵, 这种做法破坏了数据的结构, 而 TRPCA 无需进行这样的操作。对于彩色图像, RPCA 也需要将各个通道都拉成向量然后拼成矩阵或者将彩色图像变成灰度图像, 同样破坏了数据的原有结构。TRPCA 的模型为

$$\min_{\mathcal{L}, \mathcal{E}} \|\mathcal{L}\|_* + \lambda \|\mathcal{E}\|_1, \text{ s.t. } \mathcal{X} = \mathcal{L} + \mathcal{E},$$

其中, $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 表示受噪声干扰的张量, $\mathcal{L} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 和 $\mathcal{E} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 分别表示低秩张量和稀疏张量。TRPCA 的文章中严格推导出了一个新的张量核范数和其他一些相关的张量概念, 使它们具有与矩阵情况相同的联系, 将 RPCA 的模型、优化方法和理论分析技术从矩阵情况扩展到了张量情况, 证明了在一定条件下, TRPCA 模型可以从受噪声干扰的张量 \mathcal{X} 中来恢复低秩分量 \mathcal{L} 和稀疏分量 \mathcal{E} , 避免了 RPCA 在处理张量数据时的信息丢失, 在视频的恢复[10], 视频去噪[11]等实际问题有着广泛的应用。

但是该方法也存在一定的局限性, 首先, TRPCA 没有考虑彩色图像倾斜的情况, 改进这一点需要我们对 TILT 进行推广, 构造一个使用张量核范数的模型, 其次, 模型中仅仅考虑了稀疏噪声, 并没有考虑高斯噪声, 为此, 我们提出 RTILT 来克服相关工作的局限性, 我们的方法将考虑变换后张量的低秩性, 同时引入张量的 F 范数来处理高斯噪声。

2. 相关工作

2.1. 符号说明

在本文中, 我们用粗体欧拉字母表示张量, 例如 \mathcal{A} , 矩阵用大写字母来表示, 如 A ; 对于一个三阶张量 $\mathcal{A} \in \mathbb{C}^{n_1 \times n_2 \times n_3}$, 我们将它的第 (i, j, k) 个元素表示为 \mathcal{A}_{ijk} 或 a_{ijk} , 并使用 Matlab 中的形式 $\mathcal{A}(:, :, i)$ 来表示张量 \mathcal{A} 的第 i 个正面切片, 也可以简化写成 $A^{(i)}$ 。我们还用 $\|\mathcal{A}\|_1 = \sum_{ijk} |a_{ijk}|$ 代表张量 1 范数,

$\|\mathcal{A}\|_F^2 = \sqrt{\sum_{ijk} |a_{ijk}|^2}$ 代表张量 F 范数, $\langle \mathcal{A}, \mathcal{B} \rangle$ 表示两个张量 $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{n_1 \times n_2 \times n_3}$ 的内积, 定义为

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i=1}^{n_3} \langle A^{(i)}, B^{(i)} \rangle.$$

2.2. 鲁棒张量的主成分分析

鲁棒张量的主成分分析(TRPCA)是处理图像恢复和背景建模问题的常用模型, 受到最近提出的张量 - 张量积[12]的启发, TRPCA 严格定义了张量核范数, 使它的性质和关系与矩阵的情况一致。

定义 2.2.1 (张量核范数[9]) 若 $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^*$ 是 $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 的 t -SVD, 则 \mathcal{A} 的张量核范数定义为

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

其中 $r = \#\{i, \mathcal{S}(i, i, :) \neq 0\}$, \mathcal{V}^* 是 \mathcal{V} 的共轭转置, $*$ 表示张量 - 张量积, 由矩阵 - 矩阵积推广而来[12]。

鲁棒张量的主成分分析主要通过凸优化

$$\min_{\mathcal{L}, \mathcal{E}} \|\mathcal{L}\|_* + \lambda \|\mathcal{E}\|_1, \text{ s.t. } \mathcal{X} = \mathcal{L} + \mathcal{E}, \quad (1)$$

从受噪声干扰的张量 $\mathcal{X} = \mathcal{L} + \mathcal{E} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 中来恢复低秩分量 \mathcal{L} 和稀疏分量 \mathcal{E} , 其中, $\lambda = 1/\sqrt{\max(n_1, n_2)n_3}$ 。对于问题(1), 可以采用交替方向迭代法(ADMM) [13]来进行求解, 实现目标与背景的精确分离, 验证了 TRPCA 在图像恢复和背景建模问题上的有效性。

3. 我们的方法

3.1. 模型介绍

TRPCA 模型仅仅考虑了本身具有低秩性质的张量数据, 但是现实世界中的数据并不都是如此, 例如我们拍摄的彩色图像往往带有一定的倾斜角度, 在这种情况下, 该图像经过一个变换后才可能具有低秩性。由此, 本文将 TRPCA 进行推广, 通过考虑变换后张量的低秩性和稀疏性进行建模, 此外, 针对实际问题, 我们还加入了张量的 F 范数来处理高斯噪声。

关于变换的低秩图像, 我们假设一个低秩纹理 \mathcal{I}^0 位于某个场景中的一个平面上, 我们从每个某个角度拍摄到的图像 \mathcal{I} 经过变换 τ 才能的到原始低秩纹理 \mathcal{I}^0 , 即 $\mathcal{I} = \mathcal{I}^0 \circ \tau^{-1}$, $\tau: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 。除了变换外, 实际生活中的图像还可能会被像素、噪声或者遮挡所破坏, 在本文中, 我们假设只有一小部分图像像素被较大的误差所破坏, 所以, 我们可以对以下偏差进行建模, $\mathcal{I} = \mathcal{I}^0 + \mathcal{E} + \mathcal{F}$, 其中, \mathcal{E} , \mathcal{F} 是稀疏张量。因此, 我们可以得到如下模型

$$\begin{aligned} \min \|\mathcal{I}^0\|_* + \lambda_1 \|\mathcal{E}\|_1 + \lambda_2 \|\mathcal{F}\|_F^2, \\ \text{s.t. } \mathcal{I} \circ \tau + \nabla \mathcal{I} \Delta \tau = \mathcal{I}^0 + \mathcal{E} + \mathcal{F}, \\ S \Delta \tau = 0 \end{aligned} \quad (2)$$

其中 $\nabla \mathcal{I}$ 是一个三阶张量, 它的每个正面切片都是关于变换 τ 参数的导数的雅克比矩阵, S 是 τ 的线性化约束, λ_1, λ_2 是加权参数。

3.2. 模型求解

我们采用 ADMM 求解模型, 则上式的增广拉格朗日函数为

$$\begin{aligned} L_\mu(\mathcal{I}^0, \mathcal{E}, \mathcal{F}, \Delta \tau, \mathcal{Y}, Z) = \|\mathcal{I}^0\|_* + \lambda_1 \|\mathcal{E}\|_1 + \lambda_2 \|\mathcal{F}\|_F^2 + \langle \mathcal{Y}, \mathcal{I} \circ \tau + \nabla \mathcal{I} \Delta \tau - \mathcal{I}^0 - \mathcal{E} - \mathcal{F} \rangle \\ + \frac{\mu}{2} \|\mathcal{I} \circ \tau + \nabla \mathcal{I} \Delta \tau - \mathcal{I}^0 - \mathcal{E} - \mathcal{F}\|_F^2 + \langle Z, S \Delta \tau \rangle + \frac{\mu}{2} \|S \Delta \tau\|_F^2, \end{aligned} \quad (3)$$

其中, $\mu > 0$, \mathcal{Y}, Z 是拉格朗日乘子。

下面对 \mathcal{I}^0 、 \mathcal{E} 、 \mathcal{F} 和 $\Delta \tau$ 依次最小化。

1) 关于 \mathcal{I}_{k+1}^0 :

$$\mathcal{I}_{k+1}^0 = \arg \min_{\mathcal{I}^0} \frac{1}{\mu_k} \|\mathcal{I}^0\|_* + \frac{1}{2} \left\| \mathcal{I}^0 - \left(\mathcal{I} \circ \tau + \nabla \mathcal{I} \Delta \tau_k - \mathcal{E}_k - \mathcal{F}_k + \frac{\mathcal{Y}_k}{\mu_k} \right) \right\|_F^2; \quad (4)$$

2) 关于 \mathcal{E}_{k+1} :

$$\mathcal{E}_{k+1} = \arg \min_{\mathcal{E}} \frac{\lambda_1}{\mu_k} \|\mathcal{E}\|_1 + \frac{1}{2} \left\| \mathcal{I}^0 - \left(\mathcal{I} \circ \tau + \nabla \mathcal{I} \Delta \tau_k - \mathcal{I}_{k+1}^0 - \mathcal{F}_k + \frac{\mathcal{Y}_k}{\mu_k} \right) \right\|_F^2; \quad (5)$$

3) 关于 \mathcal{F}_{k+1} :

$$\mathcal{F}_{k+1} = \frac{\mu_k \left(\mathcal{I} \circ \tau + \nabla \mathcal{I} \Delta \tau_k - \mathcal{I}_{k+1}^0 - \mathcal{E}_k + \frac{\mathcal{Y}_k}{\mu_k} \right)}{2\lambda_2 + \mu_k}; \quad (6)$$

4) 关于 $\Delta \tau_{k+1}$:

$$\Delta \tau_{k+1} = \arg \min_{\Delta \tau} \frac{\mu_k}{2} \left\| \mathcal{I} \circ \tau + \nabla \mathcal{I} \Delta \tau - \mathcal{I}_{k+1}^0 - \mathcal{E}_{k+1} - \mathcal{F}_{k+1} + \frac{\mathcal{Y}_k}{\mu} \right\|_F^2 + \frac{\mu_k}{2} \left\| S \Delta \tau + \frac{Z_k}{\mu_k} \right\|; \quad (7)$$

求导得,

$$\tau_{k+1} = \begin{bmatrix} \nabla I^{(1)} \\ \nabla I^{(2)} \\ \nabla I^{(3)} \\ S \end{bmatrix}^\dagger \begin{bmatrix} -\mathcal{I} \circ \tau^{(1)} + I_{k+1}^{0(1)} + E_{k+1}^{(1)} + F_{k+1}^{(1)} - \frac{Y_k^{(1)}}{\mu_k} \\ -\mathcal{I} \circ \tau^{(2)} + I_{k+1}^{0(2)} + E_{k+1}^{(2)} + F_{k+1}^{(2)} - \frac{Y_k^{(2)}}{\mu_k} \\ -\mathcal{I} \circ \tau^{(3)} + I_{k+1}^{0(3)} + E_{k+1}^{(3)} + F_{k+1}^{(3)} - \frac{Y_k^{(3)}}{\mu_k} \\ -\frac{Z_k}{\mu_k} \end{bmatrix} \quad (8)$$

其中 \dagger 表示摩尔-彭若斯广义逆。

最后, 文章的算法总结如下:

算法 1 RTILT 算法

输入: 输入彩色图像 $\mathcal{I} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, 初始变换 τ 和权重 λ_1, λ_2 。

输出: \mathcal{I}^0 、 \mathcal{E} 、 \mathcal{F} 和 τ 。

初始化: $\varepsilon = 10^{-4}$, $\rho = 1.25$, $f = 0$ 。

1. 归一化 $\mathcal{I} \circ \tau$, 并计算 $\nabla \mathcal{I}$ 。

$$\mathcal{I} \circ \tau \leftarrow \frac{\mathcal{I} \circ \tau}{\|\mathcal{I} \circ \tau\|_F}; \quad \nabla \mathcal{I} \leftarrow \frac{\partial}{\partial \xi} \left(\frac{\text{vec}(\mathcal{I} \circ \xi)}{\|\text{vec}(\mathcal{I} \circ \xi)\|_F^2} \right) \Bigg|_{\xi=\tau};$$

其中, $\text{vec}(\cdot)$ 表示将张量的每个正面切片按列进行向量化, 再重新拼凑成矩阵。

2. (内循环)

① 固定其他参数分别按照式(4) (5) (6) (8)来更新 $\mathcal{I}_{k+1}^0, \mathcal{E}_{k+1}, \mathcal{F}_{k+1}, \Delta \tau_{k+1}$ 。

② 更新乘子 $\mathcal{Y}_{k+1}, Z_{k+1}$

$$\begin{aligned} \mathcal{Y}_{k+1} &= \mathcal{Y}_k + \mu_k (\mathcal{I} \circ \tau + \nabla \mathcal{I} \Delta \tau_{k+1} - \mathcal{I}_{k+1}^0 - \mathcal{E}_{k+1} - \mathcal{F}_{k+1}), \\ Z_{k+1} &= Z_k + \mu_k (S \Delta \tau_{k+1}) \end{aligned}$$

③ 更新 $\mu_{k+1} = \rho \mu_k$ 。

④ 检查收敛条件: $\|\mathcal{I} \circ \tau + \nabla \mathcal{I} \Delta \tau_{k+1} - \mathcal{I}_{k+1}^0 - \mathcal{E}_{k+1} - \mathcal{F}_{k+1}\|_F + \|S \Delta \tau_{k+1}\|_F < \varepsilon$ 。若不满足收敛条件, 重复步骤①~④, 若满足收敛条件, 内循环结束。

3. 更新变换: $\tau \leftarrow \tau + \Delta \tau$ 。

4. 检查收敛条件: $\|\mathcal{I}^0\|_* + \lambda_1 \|\mathcal{E}\| + \lambda_2 \|\mathcal{F}\|_F^2 - f < \varepsilon$, 若不满足收敛条件, 更新 $f = \|\mathcal{I}^0\|_* + \lambda_1 \|\mathcal{E}\| + \lambda_2 \|\mathcal{F}\|_F^2$ 再重复 1~3。若满足收敛条件, 结束。

4. 实验

本章我们将通过数值实验主要和 TRPCA 进行比较来验证我们所做工作的必要性, 我们将进行三种实验, 包括规则倾斜图像的校正问题, 彩色图像恢复问题和灰度视频的前背景分割问题。与此同时, 我们还分别针对不同的图像调整参数, 为使每种方法在每个图像下都能得到最优的恢复效果。

4.1. 规则倾斜图像校正

本节主要针对纹理结构比较规则的自然倾斜图像进行实验分析, 并与 TRPCA 方法进行比较。我们所选取的实验图像都是实际生活中常见物体的照片, 比如建筑物、黑板字、车牌、路标等。如图 1 所示, (a)为实验输入的图像, 用来选取待分析图像中的主要纹理部分, (b)中红色框是实验时选取的区域, 算法将对这部分的数据进行分析处理, (c)为应用 TRPCA 算法的实验结果图。(d)为应用本文算法的实验结果图, 即(b)中红框区域内校正的低秩图像。显然, 我们的方法可以校正自然倾斜图像, TRPCA 效果不佳。



Figure 1. Comparison of correction effects of some naturally inclined images under different algorithms

图 1. 部分自然倾斜图像在不同算法下校正效果对比

下面我们在彩色格子图像的校正问题上验证我们方法的有效性, 图 2(b)中展示了三种彩色格子图像种需要校正的部分, 由于图像有一些倾斜, 容易发现这些彩色格子图像不直接具备低秩性, 因此 TRPCA 难于处理这样的图像, 但是格子纹理具备一定的规律, 所以这些图像经过一个变换后会具备低秩性, 从图 2(d)中可以看出我们的方法可以将格子图像进行校正, TILT 方法也是曾经提出了的能够校正图像的方法, 但是该方法仅仅利用矩阵的低秩性进行建模而不是使用张量的低秩性, 在处理灰度图像时效果很好。但是对于这种彩色图像, 由于算法要求输入必须是矩阵, 需要将彩色图像转换成灰度图像进行输入或者只利用彩色图像的某一个通道进行输入, 这样的做法通常会导致原始数据的信息丢失。我们选择了将彩色图像变成灰度图像的策略, 其实验结果如图 2(c)所示, 我们发现 TILT 无法准确校正这些彩色格子图像, 原因是这些彩色格子图像变成灰度图像后格子纹理不够明显。

4.2. 彩色图像的恢复

本节中, 我们选取了来自伯克利分割数据集[14]的部分彩色图像来进行测试, 对于每幅图像, 我们随机添加了不同程度的稀疏噪声和高斯噪声, 实验结果如图 3 所示, 其中, (a)为实验输入的原始图像, (b)为随机添加噪声损坏的图像, 算法将对这部分的数据进行分析处理, (c)为应用 TRPCA 算法的实验结果

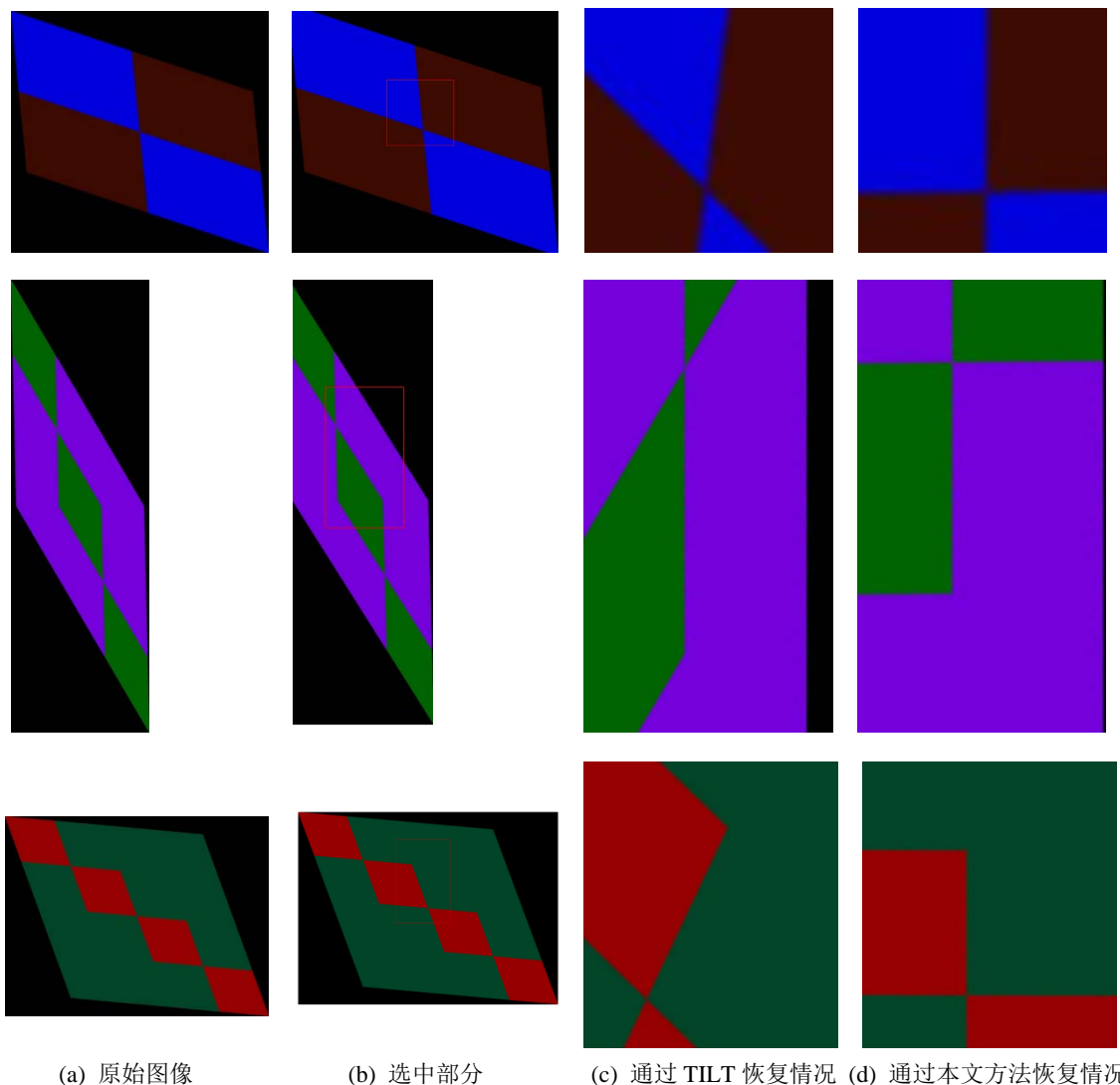
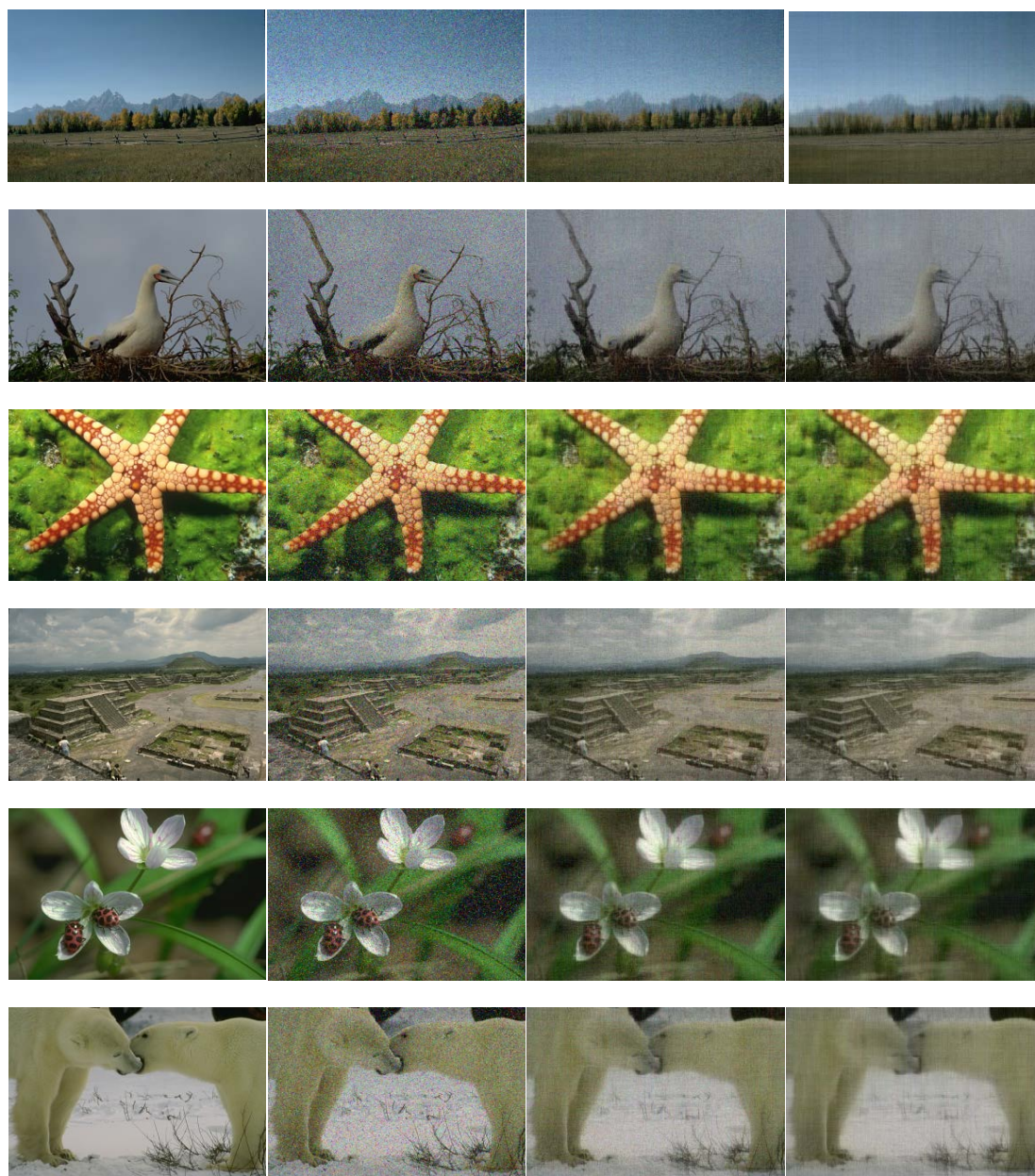


Figure 2. Comparison of correction effects of some color oblique grid images under different algorithms
图 2. 部分彩色倾斜格子图像在不同算法下校正效果对比

图, (d)为应用本文算法的实验结果图。从视角效果上看, 我们的模型在样本 1 和样本 6 上表现更佳, (d)中噪声明显少于(c), 而在其他几个样本上, 我们的方法和 TRPCA 一样有效。为了更直观的比较出两种方法的优劣, 我们还采用峰值信噪比(PSNR)作为去噪性能评价指标, PSNR 值越高, 图像的去噪性能越好, 所提算法与 TRPCA 在不同测试图像下的 PSNR 值如表 1 所示。总的来说, 我们提出的模型去噪效果更佳。

Table 1. Comparison of PSNR values on the above 6 images
表 1. 上述 6 张图像上的 PSNR 值的比较

	1	2	3	4	5	6
TRPCA	24.2301	21.6581	22.3045	22.8979	23.8765	22.8086
我们	26.4251	22.1751	22.8091	23.6925	25.4740	24.3571



(a) 原始图像 (b) 被损坏的图像 (c) TRPCA (d) 我们的方法

Figure 3. Comparison of restoration performance of color images with randomly added sparse noise and Gaussian noise

图 3. 随机添加稀疏噪声和高斯噪声的彩色图片的恢复性能比较

4.3. 灰度视频的前背景分割

在背景建模方面,水面的波纹,摇晃的树枝,飘扬的旗帜等动态背景会产生很多躁动点,以至于部分噪声分割到前背景中。因此,在本节中,我们将考虑存在高斯噪声的灰度视频前背景分割的问题。在此,我们考虑了 Campus、hall、Lobby 和 ShoppingMall 四个视频,对于每个视频,我们将视频的每一帧都随机添加高斯噪声,再分别输入给我们的模型和 TRPCA。图 4 显示了我们模型和 TRPCA 的前背景分割结果,其中(a)表示原始帧,(b)代表随机添加高斯噪声后的损坏帧、(c)和(d)分别表示由 TRPCA 和我们

的模型分割的低秩分量和稀疏分量。从这些结果中我们可以明显观察到, (d)中噪声明显少于(c), 也就是说我们的模型在分割前背景方面比 TRPCA 表现得好多, 尤其是在 Campus 和 ShoppingMall 上, 原因在于我们模型中加入了 F 范数, 可以更好的处理高斯噪声, 从而更好的分割前背景。



Figure 4. Comparison of foreground and background segmentation results for grayscale videos

图 4. 灰度视频的前背景分割结果比较

5. 结论

本文提出了一种基于鲁棒张量主成分分析的变换后低秩彩色图像的校正方法, 利用变换后图像具备的低秩性进行建模, 构造了一个使用张量核范数的模型。与 TRPCA 相比, 本文所提出的算法可以有效地校正倾斜的彩色图像, 同时本文的方法还加入了张量的 F 范数, 可以更好地处理高斯噪声和分割视频的动态背景。最后, 各种彩色图像和灰度视频上的大量实验结果证明了本文方法的有效性。

基金项目

国家自然科学基金项目(62076115)。

参考文献

- [1] Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., *et al.* (2009) Robust Face Recognition via Sparse Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **31**, 210-227. <https://doi.org/10.1109/TPAMI.2008.79>
- [2] Elad, M., Figueiredo, M.A.T. and Ma, Y. (2010) On the Role of Sparse and Redundant Representations in Image Processing. *Proceedings of the IEEE*, **98**, 972-982. <https://doi.org/10.1109/JPROC.2009.2037655>
- [3] Wright, J., Ma, Y., Mairal, J., *et al.* (2009) Sparse Representation for Computer Vision and Pattern Recognition. *Proceedings of the IEEE*, **98**, 1031-1044. <https://doi.org/10.1109/JPROC.2010.2044470>
- [4] Candès, E.J., Li, X.D., Ma, Y., *et al.* (2011) Robust Principal Component Analysis? *Journal of the ACM*, **58**, 1-37. <https://doi.org/10.1145/1970392.1970395>
- [5] Liu, G.C., Lin, Z.C. and Yu, Y. (2010) Robust Subspace Segmentation by Low-Rank Representation. *Proceedings of the 27th International Conference on Machine Learning*, Haifa, 21-24 June 2010, 663-670.
- [6] Bouwmans, T. and Zahzah, E.H. (2014) Robust PCA via Principal Component Pursuit: A Review for a Comparative Evaluation in Video Surveillance. *Computer Vision and Image Understanding*, **122**, 22-34. <https://doi.org/10.1016/j.cviu.2013.11.009>
- [7] Zhang, Z.D., Ganesh, A., Liang, X., *et al.* (2012) TILT: Transform Invariant Low-Rank Textures. *International Journal of Computer Vision*, **99**, 1-24. <https://doi.org/10.1007/s11263-012-0515-x>
- [8] Kolda, T.G. and Bader, B.W. (2009) Linear Algebra and Its Applications. *SIAM Review*, **51**, 455-500. <https://doi.org/10.1137/07070111X>
- [9] Liu, C.Y., Feng, J.S., Chen, Y.D., *et al.* (2020) Tensor Robust Principal Component Analysis with a New Tensor Nuclear Norm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**, 925-938. <https://doi.org/10.1109/TPAMI.2019.2891760>
- [10] Ji, H., Huang, S., Shen, Z., *et al.* (2011) Robust Video Restoration by Joint Sparse and Lowrank Matrix Approximation. *SIAM Journal on Imaging Sciences*, **4**, 1122-1142. <https://doi.org/10.1137/100817206>
- [11] Liu, J., Musialski, P., Wonka, P., *et al.* (2012) Tensor Completion for Estimating Missing Values in Visual Data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **35**, 208-220. <https://doi.org/10.1109/TPAMI.2012.39>
- [12] Kilmer, M.E. and Martin, C.D. (2011) Factorization Strategies for Third Order Tensors. *Linear Algebra and Its Applications*, **435**, 641-658. <https://doi.org/10.1016/j.laa.2010.09.020>
- [13] Lu, C.Y., Feng, J.S., Yan, S.C., *et al.* (2018) A Unified Alternating Direction Method of Multipliers by Majorization Minimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **40**, 527-541. <https://doi.org/10.1109/TPAMI.2017.2689021>
- [14] Martin, D., Fowlkes, C., Tal, D., *et al.* (2010) Database of Human Segmented Natural Images and Its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. *Proceedings Eighth IEEE International Conference on Computer Vision*, Vancouver, 7-14 July 2001, 416-423.