

基于低秩高阶张量逼近的图像视频恢复

王智豪¹, 刘彦²

¹西南大学, 数学与统计学院, 重庆

²重庆市市场监督管理档案信息中心, 重庆

收稿日期: 2022年2月15日; 录用日期: 2022年4月13日; 发布日期: 2022年4月20日

摘要

图像视频恢复是计算机视觉中一项基本但关键的任务, 近年来得到了广泛的研究。然而, 现有的方法存在着不可避免的缺点: 有些需要预定义秩, 有些则无法处理高阶数据。为了克服这些缺点, 本文利用图像视频数据通常具有的低秩性, 采用低秩高阶张量逼近方法实现在混合噪音的环境下的彩色视频恢复。首先, 本文建立了一个高阶张量代数框架。基于该框架, 通过设计近端算子, 提出了一种新的低秩高阶张量逼近(LRHA)方法, 旨在从被高度污染的阶张量数据中恢复出潜在的低秩部分, 从而完成图像视频恢复任务。设计了相应的算法, 并且针对多项图像视频恢复任务的实验结果表明, LRHA方法在处理相应问题方面具有优越性。

关键词

图像视频恢复, 低秩高阶张量逼近, 核范数, 张量奇异值分解

Image and Video Recovery Based on Low-Rank High-Order Tensor Approximation

Zhihao Wang¹, Yan Liu²

¹School of Mathematics and Statistics, Southwest University, Chongqing

²Market Supervision and Administration Archives Information Center, Chongqing

Received: Feb. 15th, 2022; accepted: Apr. 13th, 2022; published: Apr. 20th, 2022

Abstract

Image and video recovery is a basic but key task in computer vision and has been widely studied

in recent years. However, existing methods have inevitable disadvantages: some require pre-defined rank, while others cannot handle high-order data. In order to overcome these shortcomings, this paper uses the low-rank high-order tensor approximation method to realize color video recovery in mixed noise environment. Firstly, a high-order tensor algebraic framework is established. Based on the framework, a new low-rank high-order tensor approximation (LRHA) method is proposed by designing a proximal operator to recover potential low-rank parts from highly contaminated tensor data, so as to complete the task of image and video restoration. And the corresponding algorithm was designed. Experimental results of image and video restoration tasks show that LRHA has advantages in dealing with corresponding problems.

Keywords

Image and Video Recovery, Low-Rank High-Order Tensor Approximation, Nuclear Norm, Tensor Singular Value Decomposition

Copyright © 2022 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

随着数据科学和计算机硬件能力的发展,人们处理大数据量的高维张量成为可能。张量是多维数据建模的强大工具,如彩色图像、视频、高光图像等。以图像和视频数据为代表的高阶张量数据在传输过程中,数据会不可避免地受到各种噪声环境的影响,例如:椒盐噪声、高斯噪声、泊松噪声等,从被污染的图像视频中恢复干净数据成为图像视频处理的一个基本和必要的步骤。通常的恢复处理过程利用了数据固有的先验特征,如低秩性、稀疏性和局部相似性[1] [2] [3]。其中基于低秩的方法,例如低秩矩阵逼近(LRMA)和低秩张量逼近(LRTA) [4] [5],近年已经得到广泛的研究。

主成分分析(PCA)方法是处理恢复问题的一种基本的低秩矩阵逼近方法,为了提高PCA模型的鲁棒性,Candes提出了鲁棒主成分分析(RPCA)方法[6]。然而,低秩矩阵逼近的主要问题是仅限于处理矩阵数据,无法恢复或者要经过特殊处理才能恢复张量数据。将基于矩阵的方法拓展为3阶的张量情况需要能够处理张量的基本数学工具,这里与矩阵的情况不同,张量秩的定义到目前为止不是唯一的。至今,研究人员已经给出了多种张量秩定义以及对应的凸松弛。其中,CP秩[7]依赖于CP分解,它的值是张量分解后秩一张量的具体数目。然而,CP秩的精确计算已被证明是NP-hard。Tucker秩[8]受到更广泛的应用,并且Liu等提出了SNN方法作为Tucker秩的凸替代,但是SNN仍然不是Tucker秩的紧凸替代。除此之外,还有其他的张量秩定义,比如张量环秩、张量链秩等,可以参考文献[9] [10] [11]。特别的,Kilmer等[12]将张量作为一个整体来研究张量的奇异值分解(t-SVD)。在此基础上,Lu等[13]提出了张量主成分分析(TRPCA)方法使得我们可以在特定的条件下从观察到的三阶张量 $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ 中恢复低秩部分 \mathcal{L} 和稀疏部分 \mathcal{E} 。TRPCA和这一类的基于核范数的低秩张量逼近方法[14] [15] [16]能在保持数据结构完整的前提下有效地解决实际问题。

虽然这些方法普遍取得了良好的成果,但仍有两个主要挑战:1)一些现有的低秩逼近方法可能需要预先定义秩的值;2)这些方法主要处理矩阵和3阶张量数据,虽然一些低秩张量逼近方法可以通过重塑数据来处理高阶数据,但是这种处理会破坏数据的全局空间结构。鉴于上述缺点,本文利用[17]中提出的启发式的处理高阶张量的循环思想,设计了高阶张量的代数框架和相应的高阶张量SVD分解方法。同时,

定义了一个新的张量正向秩以及高阶张量核范数。最后基于上述要素, 本文设计出全新的低秩高阶张量逼近方法(LRHA), 它能够从高阶张量数据(阶数 $p \geq 3$)中恢复出潜在的低秩部分。具体表达形式为: $\mathcal{X}^p = \mathcal{L}^p + \mathcal{E}^p + \mathcal{G}^p$, 其中 \mathcal{L}^p 是目标低秩张量, \mathcal{E}^p 为稀疏误差项, \mathcal{G}^p 为高斯噪声项。

本工作的贡献总结如下: 1) 推导出高阶张量的共轭性质, 并将其应用于张量积和张量奇异值分解运算, 这种改进相对于不加共轭的原方法将节省一定计算成本; 2) 利用提出的基于张量正向秩的张量核范数, 成功地完善了 LRHA 所需的代数框架并给出了完整算法; 3) 通过对多项图像视频的恢复任务验证了 LRHA 方法的恢复能力, 理论上可以从任意维数的观测数据中恢复低正向秩部分。

2. 符号和数学框架

2.1. 符号

本节介绍了文中使用的一些符号。实数域和复数域分别记为 \mathbb{R} 和 \mathbb{C} 。本文用黑体欧拉字母表示张量, 上标表示张量的相应维数, 例如一个 p 阶张量可以记为 $\mathcal{A}^p \in \mathbb{C}^{n_1 \times n_2 \times \dots \times n_p}$, 矩阵用黑体大写字母表示, 例如 \mathbf{A} , 向量以黑体小写字母表示, 例如 \mathbf{a} , 用小写字母表示标量, 例如 a 。已知一个阶 3 张量 $\mathcal{A}^3 \in \mathbb{C}^{n_1 \times n_2 \times n_3}$ 可以看成是矩阵的组合, 称之为切片。对于任意 p 阶张量 $\mathcal{A}^p \in \mathbb{C}^{n_1 \times n_2 \times \dots \times n_p}$ 也可以看成是低维张量的组合, 本文称之为子张量, 它通过固定原始张量的部分维数而得到。由于本文对高阶张量的处理一般从最后一个维度逆向一直到倒数第二个维度为止, 将 $\mathcal{A}^{p-i(i_1, \dots, i_{p-i+1}, \dots, i_p)}$ 简写为 $\mathcal{A}^{p-i(i_{p-i+1}, \dots, i_p)}$, $i = 1: p-2$ 。特别的, \mathcal{A}^p 的正向切片定义为 $\mathcal{A}^{2(i_3, \dots, i_p)}$, 它是一个二阶子张量并且可以被记为 $\mathbf{A}^{(i_3, \dots, i_p)}$ 。并且本文称大小为 $1 \times 1 \times n_3 \times \dots \times n_p$ 的子张量 $\mathcal{A}^{p-2(i, j, \dots, i)}$ 为管子张量。对于任意 $\mathcal{A}^p \in \mathbb{C}^{n_1 \times n_2 \times \dots \times n_p}$, 它的复共轭定义为 $\text{conj}(\mathcal{A}^p)$, 即取 \mathcal{A}^p 每一个元素的复共轭。本文定义 p 阶张量的 l_1 -范数为 $\|\mathcal{A}^p\|_1 = \sum_{i_1, i_2, \dots, i_p} |a_{i_1, i_2, \dots, i_p}|$, 无穷范数定义为

$$\|\mathcal{A}^p\|_\infty = \max_{i_1, i_2, \dots, i_p} |a_{i_1, i_2, \dots, i_p}|。$$

2.2. DFT 共轭性质

离散傅里叶变换(DFT)在本文算法运算时起着核心作用。已知对 $\mathbf{a} \in \mathbb{R}^n$ 的傅里叶变换后结果为 $\bar{\mathbf{a}}$, 记为 $\bar{\mathbf{a}} = \mathbf{F}_n \mathbf{a} \in \mathbb{C}^n$, 其中 \mathbf{F}_n 是 DFT 矩阵, 已知 \mathbf{a} 的循环矩阵可以用 DFT 矩阵对角化,

$$\mathbf{F}_n \cdot \text{circ}(\mathbf{a}) \cdot \mathbf{F}_n^{-1} = \text{diag}(\bar{\mathbf{a}})$$

其中 $\text{diag}(\bar{\mathbf{a}})$ 表示一个对角矩阵, 其第 i 个对角项为 \bar{a}_i 。

引理 2.1 对于 $\mathcal{A}^p \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_p}$, 张量 $\bar{\mathcal{A}}^p \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_p}$, $\bar{\mathcal{A}}^p = \text{fft}(\mathcal{A}^p, [], i)$, i 从 3 到 p 。
 $\bar{\mathcal{A}}^p \in \mathbb{C}^{n_1 n_3 \dots n_p \times n_2 n_3 \dots n_p}$ 是一个块对角矩阵, 它的第 (i_3, i_4, \dots, i_p) 对角块 $\bar{\mathcal{A}}^{(i_3, i_4, \dots, i_p)}$ 是 $\bar{\mathcal{A}}$ 的第 (i_3, i_4, \dots, i_p) 正向切片, 我们有,

$$\begin{cases} \bar{\mathcal{A}}^{(1, 1, \dots, 1)} \in \mathbb{R}^{n_1 \times n_2} \\ \text{conj}(\bar{\mathcal{A}}^{(n_3 - i_3 + 2, 1, \dots, 1)}) = \bar{\mathcal{A}}^{(i_3, 1, \dots, 1)}, \\ \text{conj}(\bar{\mathcal{A}}^{(n'_3, n_4 - i_4 + 2, 1, \dots, 1)}) = \bar{\mathcal{A}}^{(i'_3, i_4, 1, \dots, 1)}, \\ \vdots \\ \text{conj}(\bar{\mathcal{A}}^{(n'_3, n'_4, 1, \dots, n_p - i_p + 2)}) = \bar{\mathcal{A}}^{(i'_3, i'_4, \dots, i'_p)}, \end{cases} \quad (2.1)$$

$$n'_j = \begin{cases} 1, i_j = 1, \\ n_j - i_j + 2, i_j \neq 1. \end{cases}; \quad i'_k = 1, \dots, n_k; \quad i'_l = 2, \dots, \left\lfloor \frac{n_k + 1}{2} \right\rfloor; \quad l = 3, \dots, p; \quad j, k = 3, \dots, p-1。$$

上述引理的证明使用到了块循环矩阵 $\tilde{\mathbf{A}}^p \in \mathbb{R}^{n_1 n_3 n_4 \cdots n_p \times n_2 n_3 n_4 \cdots n_p}$ 可通过傅里叶变换对角化的性质, 即

$$\bar{\mathbf{A}}^p = (\mathbf{F}_{n_p} \otimes \mathbf{F}_{n_{p-1}} \otimes \cdots \otimes \mathbf{F}_{n_3} \otimes \mathbf{I}_{n_1}) \tilde{\mathbf{A}}^p (\mathbf{F}_{n_p}^* \otimes \mathbf{F}_{n_{p-1}}^* \otimes \cdots \otimes \mathbf{F}_{n_3}^* \otimes \mathbf{I}_{n_2})$$

对于 $\bar{\mathbf{A}}^{(i_3, i_4, \dots, i_p)}$ 的任意元素, 我们都能根据(2.1)的索引在 $\bar{\mathbf{A}}$ 的某一个相应的正向切片中找到它的共轭。令 $\mathbf{F} = \mathbf{F}_{n_p} \otimes \mathbf{F}_{n_{p-1}} \otimes \cdots \otimes \mathbf{F}_{n_3}$ 且 $\mathbf{F}^* = \mathbf{F}_{n_p}^* \otimes \mathbf{F}_{n_{p-1}}^* \otimes \cdots \otimes \mathbf{F}_{n_3}^*$, 上面的等式可以简记

$$\bar{\mathbf{A}}^p = (\mathbf{F} \otimes \mathbf{I}_{n_1}) \tilde{\mathbf{A}}^p (\mathbf{F}^* \otimes \mathbf{I}_{n_2}).$$
 其中

$$\tilde{\mathbf{A}}^p = \begin{bmatrix} \tilde{\mathbf{A}}^{p-1(1)} & \tilde{\mathbf{A}}^{p-1(n_p)} & \cdots & \tilde{\mathbf{A}}^{p-1(2)} \\ \tilde{\mathbf{A}}^{p-1(2)} & \tilde{\mathbf{A}}^{p-1(1)} & \cdots & \tilde{\mathbf{A}}^{p-1(3)} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{A}}^{p-1(n_p)} & \tilde{\mathbf{A}}^{p-1(n_{p-1})} & \cdots & \tilde{\mathbf{A}}^{p-1(1)} \end{bmatrix}$$

$$\tilde{\mathbf{A}}^{p-1(i_p)} \in \mathbb{R}^{n_1 n_3 n_4 \cdots n_{p-1} \times n_2 n_3 n_4 \cdots n_{p-1}}$$

且

$$\tilde{\mathbf{A}}^{p-1(i_p)} = \begin{bmatrix} \tilde{\mathbf{A}}^{p-2(1, i_p)} & \tilde{\mathbf{A}}^{p-2(n_{p-1}, i_p)} & \cdots & \tilde{\mathbf{A}}^{p-2(2, i_p)} \\ \tilde{\mathbf{A}}^{p-2(2, i_p)} & \tilde{\mathbf{A}}^{p-2(1, i_p)} & \cdots & \tilde{\mathbf{A}}^{p-2(3, i_p)} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{A}}^{p-2(n_{p-1}, i_p)} & \tilde{\mathbf{A}}^{p-2(n_{p-1}-1, i_p)} & \cdots & \tilde{\mathbf{A}}^{p-2(1, i_p)} \end{bmatrix}.$$

2.3. 代数框架

对于 $\mathcal{A}^p \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}$, 本文定义

$$\text{unfold}(\mathcal{A}^p) = \begin{bmatrix} \mathcal{A}^{p-1(1)} \\ \mathcal{A}^{p-1(2)} \\ \vdots \\ \mathcal{A}^{p-1(n_p)} \end{bmatrix}, \quad \text{fold}(\text{unfold}(\mathcal{A}^p)) = \mathcal{A}^p,$$

unfold 算子将 \mathcal{A}^p 按照第 p 维度展开为 n_p 个大小为 $n_1 \times n_2 \times \cdots \times n_{p-1}$ 的张量且 fold 是它的逆算子。

$$\text{tcirc}(\mathcal{A}^p) = \tilde{\mathbf{A}}^p, \quad \text{tunfold}(\mathcal{A}^p) = \hat{\mathbf{A}}^p = \begin{bmatrix} \hat{\mathbf{A}}^{p-1(1)} \\ \hat{\mathbf{A}}^{p-1(2)} \\ \vdots \\ \hat{\mathbf{A}}^{p-1(n_p)} \end{bmatrix}, \quad \hat{\mathbf{A}}^{p-1(i_p)} = \begin{bmatrix} \hat{\mathbf{A}}^{p-2(1, i_p)} \\ \hat{\mathbf{A}}^{p-2(2, i_p)} \\ \vdots \\ \hat{\mathbf{A}}^{p-2(n_{p-1}, i_p)} \end{bmatrix}, \dots$$

其中 $\hat{\mathbf{A}}^{2(i_3, i_4, \dots, i_p)}$ 是 $\mathbf{A}^{(i_3, i_4, \dots, i_p)}$ 。

tcirc 算子将 \mathcal{A}^p 展开为它的块循环矩阵 $\tilde{\mathbf{A}}^p$; tunfold 算子将 \mathcal{A}^p 展开为大小为 $n_1 n_3 \cdots n_p \times l$ 的矩阵且 tfold 是它的逆算子 $\text{tfold}(\text{tunfold}(\hat{\mathbf{A}}^p)) = \mathcal{A}^p$ 。

定义 2.1 令 $\mathcal{A}^p \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_p}, \mathcal{B}^p \in \mathbb{R}^{n_2 \times l \times n_3 \times \cdots \times n_p}$ 。那么张量积 $\mathcal{C}^p = \mathcal{A}^p * \mathcal{B}^p$, \mathcal{C}^p 是大小为 $n_1 \times l \times n_3 \times \cdots \times n_p$ 的 p 阶张量,

$$\mathcal{C}^p = \mathcal{A}^p * \mathcal{B}^p = \text{tfold}(\text{tcirc}(\mathcal{A}^p) \cdot \text{tunfold}(\mathcal{B}^p))$$

引理 2.1 提供了一种更有效的方法来计算高阶张量积。例如, $\mathcal{A}^p \in \mathbb{R}^{n_1 \times n_2 \times 5 \times 5 \times 5}$ 和 $\mathcal{B}^p \in \mathbb{R}^{n_2 \times l \times 5 \times 5 \times 5}$ 是两

个 5 阶张量。如果想要得到它们之间的乘积, 按标准方法需要 125 次矩阵乘积。但如果使用共轭性质, 我们只需要计算 63 次乘积。

定义 2.2 令 $\mathcal{A}^p \in \mathbb{C}^{n_1 \times n_2 \times \dots \times n_p}$, 它的共轭转置为 $\mathcal{A}^{p*} \in \mathbb{C}^{n_2 \times n_1 \times n_3 \times \dots \times n_p}$, 即

$$\mathcal{A}^{p*} = \text{fold} \left(\begin{bmatrix} \mathcal{A}^{p-1(1)*} \\ \mathcal{A}^{p-1(n_p)*} \\ \vdots \\ \mathcal{A}^{p-1(2)*} \end{bmatrix} \right)。$$

定义 2.3 令 $\mathcal{I}^p \in \mathbb{R}^{n \times n \times n_3 \times \dots \times n_p}$, 当它的第一个正向切片为大小为 $n \times n$ 的单位矩阵其他切片全为 0 时称它为单位张量。对于合适大小的 \mathcal{A}^p 有 $\mathcal{A}^p * \mathcal{I}^p = \mathcal{A}^p, \mathcal{I}^p * \mathcal{A}^p = \mathcal{A}^p$ 。

定义 2.4 令 $\mathcal{Q}^p \in \mathbb{R}^{n \times n \times n_3 \times \dots \times n_p}$, 若 $\mathcal{Q}^{p*} * \mathcal{Q}^p = \mathcal{Q}^p * \mathcal{Q}^{p*} = \mathcal{I}^p$, 便称这个实值张量是一个 p 阶正交张量。

定义 2.5 如果张量的每一个正向切片都是对角矩阵, 则该张量称为 f -对角张量。

定理 2.6 令 $\mathcal{A}^p \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_p}$ 是一个 p 阶实值张量。那么 \mathcal{A}^p 能被分解为

$$\mathcal{A}^p = \mathcal{U}^p * \mathcal{S}^p * \mathcal{V}^{p*}$$

其中 $\mathcal{U}^p \in \mathbb{R}^{n_1 \times n_1 \times n_3 \times \dots \times n_p}$, $\mathcal{V}^p \in \mathbb{R}^{n_2 \times n_2 \times n_3 \times \dots \times n_p}$ 是正交张量, 且 $\mathcal{S}^p \in \mathbb{R}^{n_1 \times n_2 \times n_3 \times \dots \times n_p}$ 是一个 f -对角张量。

已知矩阵的奇异值具有递减性质, 在 t-SVD 的处理过程中, 令 $\mathcal{A}^p = \mathcal{U}^p * \mathcal{S}^p * \mathcal{V}^{p*}$, \mathcal{S}^p 的第一个正向切片对角线上元素 $\mathcal{S}_{i,i}^{(1,\dots,1)}$ 具有相同的递减性质: $i = 1:n(2)$, $n(2) = \min(n_1, n_2)$, $\mathcal{S}_{i,i}^{(1,\dots,1)} \geq \mathcal{S}_{i+1,i+1}^{(1,\dots,1)}$ 。可以通过 DFT 的处理得到以下性质, $L_p = n_3 n_4 \dots n_p$

$$\mathcal{S}_{i,i}^{(1,\dots,1)} = \frac{1}{L_p} \sum_{i_p=1}^{n_p} \sum_{i_{p-1}=1}^{n_{p-1}} \dots \sum_{i_3=1}^{n_3} \bar{\mathcal{S}}_{i,i}^{(i_3, \dots, i_p)} \quad (2.2)$$

定义 2.7 对于 $\mathcal{A}^p \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_p}$, 张量正向秩(或张量 1, 2 维秩)为 $\text{rank}_f(\mathcal{A}^p)$, 张量正向秩的值为 \mathcal{S}^p 的非零管子张量的个数, 它们是通过 \mathcal{A}^p 进行 t-SVD 得到的, 即

$$\text{rank}_f(\mathcal{A}^p) = \#\{\mathcal{S}_{i,i}^{(\dots)} \neq 0\} \quad (2.3)$$

实际上张量正向秩等于 $\mathcal{S}^{p(i,i,1,\dots,1)}$ 非零项的个数。显然, 我们可以通过旋转张量即改变张量模式(mode)的排列顺序得到类似的秩。例如, p 阶张量 \mathcal{A}^p 的 1,4 维秩($\text{rank}_{1,4}$)是通过固定一维和四维, 然后进行 t-SVD 得到的。从图 1 可以看出, 无论 t-SVD 是基于一个相对较低的正向秩还是基于 1,4 维秩, 都可以在很小的相应秩的情况下很好地逼近原始视频, 这证明了图像视频数据通常有一定的低秩性, 从而保证低秩高阶张量逼近的合理性。



Figure 1. Color video can be approximated by a low-rank tensor
图 1. 彩色视频可以用低秩张量近似

定义 2.8 令 $\mathcal{A}^p = \mathcal{U}^p * \mathcal{S}^p * \mathcal{V}^{p*}$ 作为 $\mathcal{A}^p \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_p}$ 的 t-SVD 结果。 \mathcal{A}^p 的张量核范数定义为

$$\|\mathcal{A}^p\|_* := \sum_{i=1}^r S_{i,i}^{(1,\dots,1)} \quad (2.4)$$

其中 $r = \text{rank}_f(\mathcal{A}^p)$ 。

3. 低秩高阶张量逼近(LRHA)算法

本文最初的问题是从高度污染的测量值 $\mathcal{X}^p = \mathcal{L}^p + \mathcal{E}^p + \mathcal{G}^p$ 中恢复所需要的低秩张量 \mathcal{L}^p ， \mathcal{E}^p 为稀疏误差项而 \mathcal{G}^p 为高斯噪声项。这一节将展示了求解该问题算法的详细过程，为了公式简洁，在本节中略去张量左上角的标识。本文采用标准的交替方向乘子法(ADMM)来解决目标问题，首先将原问题具体表示为：

$$\min_{\mathcal{L}, \mathcal{E}, \mathcal{G}} \|\mathcal{L}\|_* + \lambda \|\mathcal{E}\|_1 + \alpha \|\mathcal{G}\|_F^2, \text{ s.t. } \mathcal{L} + \mathcal{S} + \mathcal{G} = \mathcal{X} \quad (3.1)$$

其中 μ, λ 是事先预定好的平衡参数。 \mathcal{L} 是潜在的低秩张量， \mathcal{E} 是稀疏误差， \mathcal{G} 是高斯噪声项。下面给出原问题的增广拉格朗日函数：

$$L_\rho(\mathcal{L}, \mathcal{E}, \mathcal{G}; \mathcal{Y}) = \|\mathcal{L}\|_* + \lambda \|\mathcal{E}\|_1 + \alpha \|\mathcal{G}\|_F^2 + \frac{\rho}{2} \left\| \mathcal{L} + \mathcal{S} + \mathcal{G} - \mathcal{X} + \frac{\mathcal{Y}}{\rho} \right\|_F^2 \quad (3.2)$$

其中 \mathcal{Y} 是拉格朗日乘子， $\rho > 0$ 是正惩罚参数。总的 ADMM 算法见算法 1。

3.1. 更新 \mathcal{L}

通过固定其他变量，更新 \mathcal{L} 的子问题可以写作：

$$\mathcal{L}_{k+1} = \arg \min_{\mathcal{L}} \frac{1}{\rho_k} \|\mathcal{L}\|_* + \frac{1}{2} \|\mathcal{L} - \mathcal{A}_k\|_F^2 \quad (3.3)$$

其中 $\mathcal{A}_k = (\rho_k(\mathcal{X} - \mathcal{E}_k - \mathcal{G}_k) - \mathcal{Y}_k) / 2\rho_k$ 。在三阶张量的情形下，可以将问题转化到矩阵奇异值阈值(SVT)解决。类似的，本文引入近端算子，从而使得上述问题等价于下式：

$$\min_{\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_p}} \tau \|\mathcal{X}\|_* + \frac{1}{2} \|\mathcal{X} - \mathcal{Y}\|_F^2 \quad (3.4)$$

根据高阶张量的分解过程，可以根据文献[13]得到 SVT 算子作为上式的封闭解。令 $\mathcal{Y} = \mathcal{U} * \mathcal{S} * \mathcal{V}^*$ ，对于任意 $\tau > 0$ ，t-SVT 算子 $\mathcal{D}_\tau(\mathcal{Y}) := \mathcal{U} * \mathcal{S}_\tau * \mathcal{V}^*$ ，其中 $\mathcal{S}_\tau = \text{iff}(\bar{\mathcal{S}}, [\cdot], i)$ ， i 从 p 到 3。

3.2. 更新 \mathcal{E} 和 \mathcal{G}

通过固定其他变量，更新 \mathcal{E} 的子问题可以写作：

$$\mathcal{E}_{k+1} = \arg \min_{\mathcal{E}} \frac{\lambda}{\rho_k} \|\mathcal{E}\|_1 + \frac{1}{2} \|\mathcal{E} - \mathcal{B}_k\|_F^2 \quad (3.5)$$

其中 $\mathcal{B}_k = \mathcal{X} - \mathcal{L}_{k+1} - \mathcal{G}_{k+1} + \mathcal{Y}_k / \rho_k$ 。上式(3.5)的封闭解可以通过软阈值算子得到[18]，即

$$\mathcal{E}_{k+1} = \text{Soft}_{\lambda/\rho_k}(\mathcal{B}_k)$$

其中 $\text{Soft}_\rho(\mathbf{x}) := \text{sign}(\mathbf{x}) * \max\{|\mathbf{x}| - \rho, 0\}$ 。类似的，更新 \mathcal{G} 的子问题可以写作：

$$\mathcal{G} = \arg \min_{\mathcal{G}} \alpha \|\mathcal{G}\|_F^2 + \frac{\rho}{2} \|\mathcal{G} - \mathcal{C}\|_F^2 \quad (3.6)$$

其中 $\mathcal{C} = \mathcal{X} - \mathcal{L} - \mathcal{E} - \mathcal{Y} / \rho$ 。(3.6)的封闭解为：

$$\mathcal{G} = \rho(\mathcal{C})/(\rho + 2\alpha)$$

算法1低秩高阶张量逼近(LRHA)算法

输入: 张量数据 $\mathcal{X}^p \in \mathbb{R}^{n_1 \times n_2 \times n_3 \times \dots \times n_p}$, $\tau > 0$, 参数 λ, α

初始化: $\mathcal{L}_0^p = \mathcal{S}_0^p = \mathcal{Y}_0^p = 0, \rho = 1.5, \mu_0 = 1e-3, \mu_{\max} = 1e10, \epsilon = 1e-8$

当不收敛时, 执行

1: 通过(3.3)更新 \mathcal{L}_{k+1}^p ;

2: 通过(3.5)更新 \mathcal{E}_{k+1}^p ;

3: 通过(3.6)更新 \mathcal{G}_{k+1}^p ;

4: $\mathcal{Y}_{k+1}^p = \mathcal{Y}_k^p + \mu_k (\mathcal{L}_{k+1}^p + \mathcal{E}_{k+1}^p + \mathcal{G}_{k+1}^p - \mathcal{X}^p)$; $\mu_{k+1} = \min(\rho\mu_k, \mu_{\max})$;

5: 检查收敛条件:

$$\|\mathcal{L}_{k+1}^p - \mathcal{L}_k^p\|_{\infty} \leq \epsilon, \quad \|\mathcal{E}_{k+1}^p - \mathcal{E}_k^p\|_{\infty} \leq \epsilon, \quad \|\mathcal{G}_{k+1}^p - \mathcal{G}_k^p\|_{\infty} \leq \epsilon$$

$$\|\mathcal{L}_{k+1}^p + \mathcal{E}_{k+1}^p + \mathcal{G}_{k+1}^p - \mathcal{X}^p\|_{\infty} \leq \epsilon$$

4. 实验

在这一章中, 本文应用 LRHA 进行多种视频恢复任务来验证该方法恢复高阶张量数据的能力, 包括视频补全、视频去噪以及视频背景建模。为了实验的一致性, 本文在所有实验中设置 $\lambda = 1/\sqrt{n(1)L_p}$, $\alpha = 10$ 。在实际应用中, 对于不同的数据可以通过更仔细地调出不同的最优参数进一步提高性能。

4.1. 视频补全

在本节中, 彩色视频被用来从部分观测项中进行低秩张量补全来评估本文的方法的实用性。张量补全相当于对于只受到椒盐噪声的张量进行恢复处理。本文使用 yuv 彩色视频数据¹进行测试。对于实验中的每个视频序列, 为了控制计算成本而使用前 150 帧进行测试, 格式为网站提供的 QCIF 格式, 其中每一帧大小为 $144 \times 176 \times 3$ 。在缺失值的设置上, 对于一个大小为的彩色视频张量, 随机设置 $m = 3phwf$ 个元素来观察, 在本实验中, 考虑 $p = 0.3$, $p = 0.4$ 和 $p = 0.5$ 。同时本文应用 HaLRTC [8], TMac [19], NRTAC [20], TNN [13]这四种先进的张量修补方法进行视频恢复, 并比较它们的性能。在实验中, 这些比较方法均使用了参考文献中建议的参数设置, 其中 TNN 是将数据的前两个维度直接拼接, 再进行 3 阶的张量补全。实验中用于判断视频补全质量标准的 PSNR 值计算公式为:

$$\text{PSNR} = 10 \log_{10} \left(\frac{n_1 n_2 n_3 n_4 \|\mathcal{M}\|_{\infty}^2}{\|\hat{\mathcal{X}} - \mathcal{M}\|_F^2} \right)$$

越高的 PSNR 值证明恢复的效果越好。实验结果见表 1, 部分结果见图 2, 图 2 中第一列为各个视频的样帧, 第二列为该样帧特定采样率下的观测(比例为 $p = 0.5$), 后列分别为 HaLRTC, TMac, NRTAC, TNN 和本文的方法恢复的结果。显然, 无论是 PSNR 值还是直观的恢复效果, 都是本文的方法最优的, 证明了 LRHA 方法在张量补全上的适用性。

4.2. 视频去噪

本节应用 LRHA 对随机噪声干扰下的彩色视频图像进行恢复。考虑计算成本, 每一个视频都被预先

¹<http://trace.eas.asu.edu/yuv/>.

设置为大小为 $144 \times 176 \times 3 \times 100$ 的 4 阶张量。对于每一个视频, 本文添加了 $\kappa = 0.3$ 的脉冲噪声以及 $\sigma = 0.01$ 的高斯噪声, 同时本文应用基于 SNN [21], TNN 的两种高阶张量去噪方法进行比较, 并且添加了 LRHA 的 DCT 变体(即傅里叶变换时的 DFT 矩阵替换为 DCT 矩阵)来验证共轭方法是否能节省计算成本。在实验中, 这些比较方法均使用了参考文献中建议的参数设置, 其中基于 TNN 的方法 TRPCA 仍然是将数据的前两个维度直接拼接, 再进行 3 阶的去噪。



Figure 2. Comparison of video completion performance
图 2. 视频补全表现比较

Table 1. PSNR values of different tensor completion methods
表 1. 不同张量补全方法的 PSNR 指标

视频	缺失比例(%)	HaLRTC	TMac	NRATC	TNN	LRHA
Highway	30	30.897	30.664	34.731	33.822	36.237
	40	32.702	31.082	36.846	35.317	38.053
	50	34.389	31.462	39.085	36.910	39.844
Coastguard	30	24.655	27.486	32.229	29.619	34.009
	40	26.516	29.644	35.351	32.114	36.953
	50	28.420	31.116	38.453	34.632	39.772
Hall	30	29.021	32.404	34.303	36.031	38.547
	40	31.464	33.117	37.445	37.698	40.781
	50	33.735	33.776	40.556	39.275	42.941
Carphone	30	29.011	32.947	33.366	33.589	36.459
	40	31.445	33.779	36.236	35.831	38.879
	50	33.683	34.525	38.947	38.069	41.193
Bridge	30	30.374	30.151	33.307	36.339	38.104
	40	32.189	30.457	35.519	37.771	39.926
	50	33.923	30.745	37.775	39.204	41.604
News	30	26.997	30.732	32.329	32.437	36.349
	40	29.545	34.018	35.616	34.615	39.201
	50	31.998	35.778	38.898	36.774	41.923

图3是视频去噪结果, 第一列为视频样帧, 第二列该样帧加噪后的观测, 后列分别为 SNN, TRPCA, 本文的方法以及它的 DCT 变体恢复的结果。从实验结果表2和图3可以看出以下几点: 首先基于 Tucker 秩的 SNN 方法在时间以及 PSNR 指标上都要次于基于管秩的 TNN 方法, 其原因主要是因为 SNN 并不是目标的凸包络, 以及 Tucker 方法普遍耗时; 本文的方法与 TRPCA 相比, 时间上要稍微长一些, 原因是高阶张量的 t-SVD 需要更多的傅里叶变换过程, 但由于 LRHA 相较而言能够充分利用空间结构信息, 恢复效果普遍优于 TRPCA; 使用共轭方法的 LRHA-DFT 在恢复效果上接近于 DCT 变体, 但在时间指标上具有明显的提升。

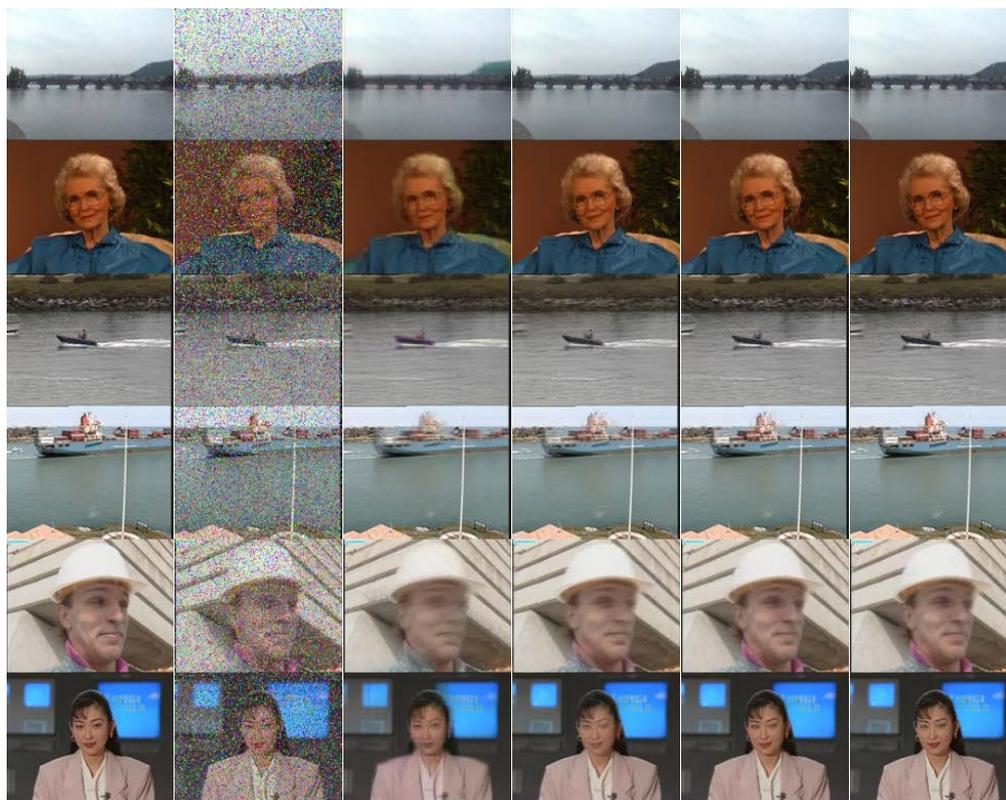


Figure 3. Comparison of video denoising performance
图3. 视频去噪表现比较

Table 2. PSNR comparison of video denoising
表2. 视频去噪 PSNR 比较

视频	SNN		TRPCA		LRHA-DFT		LRHA-DCT	
	PSNR	Time	PSNR	Time	PSNR	Time	PSNR	Time
Bridge	34.547	171.765	38.990	45.520	40.720	98.296	40.181	152.641
Grandma	27.865	154.718	36.109	42.099	38.234	97.781	38.858	152.336
Ship	26.429	158.779	27.657	38.265	31.838	99.977	30.516	151.766
Boat	30.224	152.168	34.385	39.888	40.640	103.084	41.511	152.959
Worker	26.667	153.004	29.250	39.908	31.383	97.176	31.918	153.396
Akiyo	31.494	150.758	35.589	39.490	40.641	100.38	40.152	154.304

4.3. 视频背景建模

本节应用 LRHA 对随机噪声干扰下的彩色视频图像进行背景建模处理。本节数据为 I2R 数据², 为了得到较好的背景建模效果, 每一个视频都被预先设置为大小为 $h \times w \times 3 \times 300$ 的 4 阶张量, 并添加了 $\kappa = 0.3$ 的脉冲噪声以及 $\sigma = 0.01$ 的高斯噪声, 由于处理前的前景已经受到噪声污染, 后文的前景图像是通过恢复的背景图像与干净图像直接相减得到的。图 4 展示了 LRHA 方法对不同视频进行背景建模的结果, 每一列从上到下依次为: 原始图像, 噪声图像, 前景, 背景。可以看出该方法能够在复杂噪声情况下能够去除噪声以及运动中的人物, 从而较好地完成背景建模任务。

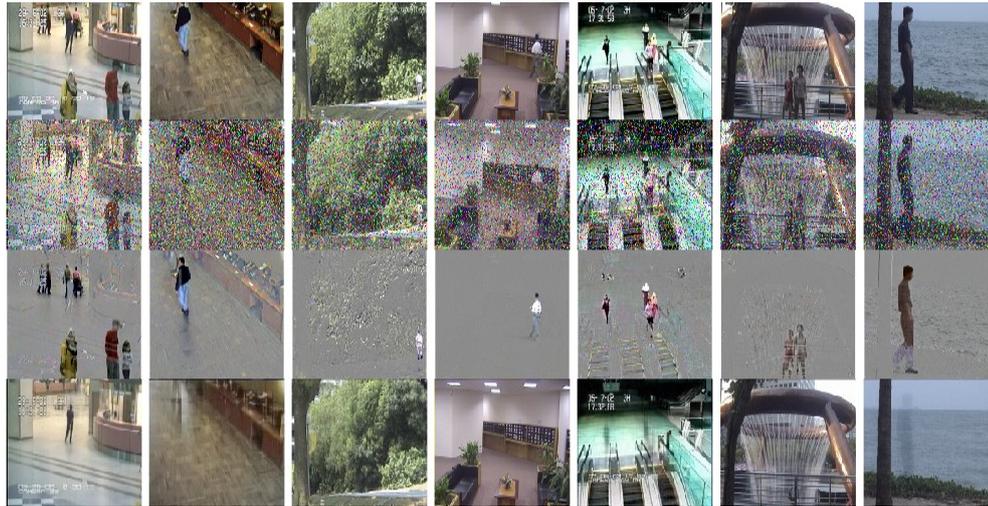


Figure 4. Background modeling results
图 4. 背景建模结果

5. 总结

本文提出了一种新的低秩高阶张量逼近方法, 利用高阶张量代数框架将原本只能作用于 2 阶或者 3 阶数据的低秩逼近方法拓展到任意维度, 并设计了完整的算法。该方法与传统低维方法的优势在于能够利用张量整体结构信息。张量补全、张量去噪、张量背景建模实验证明了本文的方法可以在采用适当参数的前提下, 有效地对以彩色视频为代表的高阶张量进行恢复处理, 为对应现实问题提供了一种实用的解决方案。

致 谢

感谢导师的精心指导以及各位同门在百忙之中提供的无私帮助, 感谢测试数据所有者的无私奉献。在论文的写作过程中引用了学界大量的研究成果, 在此谨向这些成果的作者表示衷心的感谢。

参考文献

- [1] Dong, W., Li, X., Zhang, L. and Shi, G. (2011) Sparsity-Based Image Denoising via Dictionary Learning and Structural Clustering. *Conference on Computer Vision and Pattern Recognition 2011*, Colorado Springs, 20-25 June 2011, 457-464. <https://doi.org/10.1109/CVPR.2011.5995478>
- [2] Liu, Y. and Wang, Z. (2015) Simultaneous Image Fusion and Denoising with Adaptive Sparse Representation. *IET Image Processing*, **9**, 347-357.

²<http://perception.i2r.a-star.edu.sg/bkmodel/bkindex.html>.

-
- [3] Rajpoot, N. and Butt, I. (2012) A Multiresolution Framework for Local Similarity Based Image Denoising. *Pattern Recognition*, **45**, 2938-2951. <https://doi.org/10.1016/j.patcog.2012.01.023>
- [4] Goldfarb, D. and Qin, Z. (2014) Robust Low-Rank Tensor Recovery: Models and Algorithms. *SIAM Journal on Matrix Analysis and Applications*, **35**, 225-253. <https://doi.org/10.1137/130905010>
- [5] Wright, J., Ganesh, A., Rao, S., Peng, Y. and Ma, Y. (2009) Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices via Convex Optimization. *Advances in Neural Information Processing Systems*, **22**, 2080-2088.
- [6] Candès, E.J., Li, X., Ma, Y. and Wright, J. (2011) Robust Principal Component Analysis? *Journal of the ACM*, **58**, Article No. 11. <https://doi.org/10.1145/1970392.1970395>
- [7] Kiers, H.A.L. (2000) Towards a Standardized Notation and Terminology in Multiway Analysis. *Journal of Chemometrics*, **14**, 105-122.
- [8] Liu, J., Musialski, P., Wonka, P. and Ye, J. (2013) 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>
- [9] Liu, X., Jing, X.-Y., Tang, G., Wu, F. and Dong, X. (2020) Low-Rank Tensor Completion for Visual Data Recovery via the Tensor Train Rank-1 Decomposition. *IET Image Processing*, **14**, 114-124.
- [10] Liu, Y. and Long, Z. (2019) Image Completion Using Low Tensor Tree Rank and Total Variation Minimization. *IEEE Transactions on Multimedia*, **21**, 338-350. <https://doi.org/10.1109/TMM.2018.2859026>
- [11] Oseledets, I.V. (2011) Tensor-Train Decomposition. *SIAM Journal on Scientific Computing*, **33**, 2295-2317. <https://doi.org/10.1137/090752286>
- [12] Semerci, O., Hao, N., Kilmer, M.E. and Miller, E.L. (2014) Tensor-Based Formulation and Nuclear Norm Regularization for Multienergy Computed Tomography. *IEEE Transactions on Image Processing*, **23**, 1678-1693. <https://doi.org/10.1109/TIP.2014.2305840>
- [13] Lu, C., Feng, J., Chen, Y., Liu, W., Lin, Z. and Yan, S. (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>
- [14] Feng, J., Yang, L.T., Dai, G., Wang, W. and Zou, D. (2019) A Secure High-Order Lanczos-Based Orthogonal Tensor SVD for Big Data Reduction in Cloud Environment. *IEEE Transactions on Big Data*, **5**, 355-367. <https://doi.org/10.1109/TBDATA.2018.2803841>
- [15] Song, G., Ng, M.K., and Zhang, X. (2019) Robust Tensor Completion Using Transformed Tensor SVD. arXiv preprint arXiv:1907.01113.
- [16] Zhang, X. (2019) A Nonconvex Relaxation Approach to Low-Rank Tensor Completion. *IEEE Transactions on Neural Networks and Learning Systems*, **30**, 1659-1671. <https://doi.org/10.1109/TNNLS.2018.2872583>
- [17] Martin, C.D., Shafer, R. and LaRue, B. (2013) An Order-P Tensor Factorization with Applications in Imaging. *SIAM Journal on Scientific Computing*, **35**, A474-A490. <https://doi.org/10.1137/110841229>
- [18] Tibshirani, R. (1996) Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, **58**, 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- [19] Xu, Y., Hao, R., Yin, W. and Su, Z. (2017) Parallel Matrix Factorization for Low-Rank Tensor Completion. *Inverse Problems and Imaging*, **9**, 601-624. <https://doi.org/10.3934/ipi.2015.9.601>
- [20] Zhang, Z. and Aeron, S. (2016) Exact Tensor Completion Using T-SVD. *IEEE Transactions on Signal Processing*, **65**, 1511-1526. <https://doi.org/10.1109/TSP.2016.2639466>
- [21] Huang, B., Mu, C., Goldfarb, D. and Wright, J. (2015) Provable Models for Robust Low-Rank Tensor Completion. *Pacific Journal of Optimization*, **11**, 339-364.