



## 基于张量分解的红外机场监控视频压缩技术

毛敏<sup>1</sup>, 刘宇洋<sup>1,2\*</sup>, 张昆<sup>1</sup>, 徐昱<sup>1</sup>, 孙悦<sup>1</sup>

(1. 信阳师范大学 物理电子工程学院, 信阳 464000; 2. 电子科技大学 信息与通信工程学院, 成都 611731)

**摘要:** 与传统视频相比, 大景深红外机场监控视频的背景和前景区域往往变化较平缓, 纹理信息较少, 因此其时域和空域的相关性更强。虽然可用现有视频编码工具对红外监控视频进行压缩, 但只能利用局部的视频信息去除冗余, 且编码复杂度较高。该文从数据的全局角度出发, 将整个红外监控视频视为一个三阶张量, 然后采用张量分解近似表示以降低数据量, 最后将分解得到的系数矩阵进行量化进一步提升压缩效率。同时, 采用张量的 CP 分解和 Tucker 分解, 分析对比了两种张量分解方法的压缩性能。实验结果表明, 相较于 HEVC, 对于视频存在大量平缓区域的情况, Tucker 分解可取得 77.7% 的 BD-rate 节省, 但对于内容复杂和运动较大的情况, CP 分解和 Tucker 分解的 BD-PSNR 下降超 10 dB; CP 分解和 Tucker 分解的编码时间仅为 HEVC 的 0.57% 和 2.25%。

**关键词:** 视频压缩; 红外监控视频; 张量表示; 张量分解; 量化

中图分类号: TP75

文献标志码: A

DOI: 10.12179/1672-4550.20240339

## Technology of Tensor Decomposition based Infrared Video Compression for Airport Surveillance Scenario

MAO Min<sup>1</sup>, LIU Yuyang<sup>1,2\*</sup>, ZHANG Kun<sup>1</sup>, XU Yu<sup>1</sup>, SUN Yue<sup>1</sup>

(1. College of Physics and Electronic Engineering, Xinyang Normal University, Xinyang 464000, China; 2. School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

**Abstract:** Compared with traditional videos, infrared videos in airport surveillance scenario with a large depth of field consist of smooth background and foreground areas with less textures, thus processing stronger temporal and spatial correlations. Although commonly used video coding tools can be utilized to compress infrared videos, they only utilize local information to reduce the redundancy, resulting in high computational complexity. Based on a global perspective of the data, this paper considers the whole infrared surveillance video as a third-order tensor, then adopts tensor decomposition for approximate representation to reduce the data volume, and finally quantizes the resulting factor matrices to further improve compression efficiency. Meanwhile, CP decomposition and Tucker decomposition are adopted to analyze and compare the compression performance of these two tensor decomposition methods. Experimental results show that, compared with HEVC, Tucker decomposition can achieve about 77.7% BD-rate savings for videos containing lots of smooth areas. For videos with complex content and significant motion, however, both CP decomposition and Tucker decomposition suffer from BD-PSNR loss of over 10 dB. Nevertheless, the encoding time of CP decomposition and Tucker decomposition is just about 0.57% and 2.25% of that of HEVC, respectively.

**Key words:** video compression; infrared surveillance video; tensor representation; tensor decomposition; quantization

近年来, 我国持续保持交通建设飞速发展。《交通强国建设纲要 2019》中明确指出, 我国要在 2035 年基本建成交通强国。在各类交通运输行

业中, 我国民航运输吞吐量近几年持续攀升, 民航运输总量现已排名世界第二。庞大的民航运输总量除了需要进一步提升机场运行效率, 还需加

收稿日期: 2024-06-27

基金项目: 河南省教育厅重点科研项目(23A510014); 河南省科技计划项目(242102221046); 河南省高等学校重点研究项目(22A510009)。

作者简介: 毛敏, 硕士, 助教, 主要从事图像与视频处理方面的研究。E-mail: maomin@xynu.edu.cn

\* 通信作者: 刘宇洋, 博士, 讲师, 主要从事图像与视频处理方面的研究。E-mail: yuyangliu@xynu.edu.cn

强对机场场面的监管,进而保障机场安全稳定运营。为了提升机场场面的监视能力,机场除了安置大量可见光监控摄像头外,还会安置红外摄像头以解决夜晚低光照情况下监控能力不足的问题。虽然视频监控配合目标检测、目标识别等算法可大大降低人力和物力成本,但伴随而来的海量视频数据(包括可见光监控视频和红外监控视频)为存储和传输带来巨大压力。

为了降低视频数据存储和传输成本,国际标准组织先后推出了多代视频编码标准,如高级视频编码(advanced video coding, AVC)H.264<sup>[1]</sup>、高效视频编码(high efficiency video coding, HEVC)H.265<sup>[2]</sup>、通用视频编码(versatile video coding, VVC)H.266<sup>[3]</sup>等。这些视频编码标准采用混合编码结构,包括预测、变换、量化、熵编码等模块,虽然每代视频编码标准的编码性能都比前一代高,但其编码复杂度也随之大幅增加。以高效视频编码标准 H.265/HEVC 为例,在同等视觉质量下, H.265/HEVC 可比上一代视频编码标准 H.264/AVC 节省约 50% 的码率,但其编码复杂度增加了约 5 倍。近年一些文献提出了多种降低编码复杂度的方法,虽然能够有效降低编码复杂度,但编码性能也随之降低<sup>[4-8]</sup>。

与传统彩色图像相比,红外图像纹理较少,往往包含大量平缓的区域,空域相关性更强。针对红外视频数据的高效压缩,由于可用现有视频编码工具进行压缩,因此关于其高效编码的研究相对较少<sup>[9-11]</sup>。文献[9]分别采用 JPEG2000、JPEG-Xt 和 HEVC 对红外图像进行压缩,通过对比实验结果得出, HEVC 的压缩性能最高。文献[10]针对多亮度显示红外视频(multi-intensity illuminated infrared video, MIIV)这一新型数据,将不同亮度下的红外视频视为独立视点,采用多视点加深度的编码框架进行压缩,与 HEVC 相比,在 3 个视点的条件下平均可节省 9.56% 的 BD-rate。文献[11]针对无人机红外视频数据,提出了一种基于 HEVC 的全局运动搜索算法,至少可节省 0.50% 的 BD-rate。除了采用上述编码标准压缩视频外,还有一些基于 Tucker 分解的压缩方法<sup>[12-14]</sup>,对多光谱图像/视频进行压缩,取得了较好的压缩性能。

不同于上述现有视频编码工具的压缩方法,本文提出了基于张量分解的压缩方法。首先将红

外机场监控视频数据用张量表示,然后分别采用张量的 CP(candecomp/parafac)分解和 Tucker 分解得到红外监控视频数据线性近似表达,最后将分解后的系数矩阵量化后传输至解码端。在视频包含大量平缓区域的情况下,与 HEVC 相比, Tucker 分解可节省 77.70% 的 BD-rate,但对于内容较复杂的视频, CP 分解和 Tucker 分解的压缩性能较低,在高码率情况下, BD-峰值信噪比(peak signal-to-noise ratio, PSNR)降低超 10 dB;但 CP 分解和 Tucker 分解的编码时间较少, CP 分解的编码时间仅为 HEVC 的 0.56%, Tucker 分解的编码时间为 HEVC 的 2.56%。

## 1 张量分解

### 1.1 张量表示

张量可表示多种不同维度空间的数据,包括标量、向量、矩阵及三维乃至更高维度空间的数据。如一幅尺寸为  $m \times n$  的灰度图像  $I$ ,通常可表示为矩阵  $I \in \mathbb{R}^{m \times n}$ ,其中  $m$  和  $n$  分别为图像的宽度和高度。根据张量的定义,还可将灰度图像  $I$  定义为一个二阶张量。对于只包含亮度信息的视频数据来说,需在三维数据空间中表示,可将其定义为一个三阶张量  $X \in \mathbb{R}^{W \times H \times K}$ ,如图 1 所示。其中,  $W$  和  $H$  分别表示视频的宽和高,  $K$  为视频的帧数,张量中的每个元素即为视频的像素值。

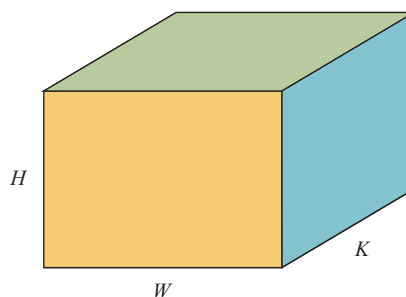


图 1 三阶张量示意图

为了便于计算,高阶张量可按照不同的模式展开为矩阵。同样以图 1 所示的三阶张量为例,可以按照 3 个不同的模式(分别为模-1、模-2 和模-3)将三阶张量展开为 3 个矩阵,如图 2 所示<sup>[15]</sup>。其中,按模-1 展开的矩阵记为  $X_{(1)} \in \mathbb{R}^{W \times HK}$ ,按模-2 展开的矩阵记为  $X_{(2)} \in \mathbb{R}^{W \times HK}$ ,按模-3 展开的矩阵记为  $X_{(3)} \in \mathbb{R}^{W \times HK}$ 。

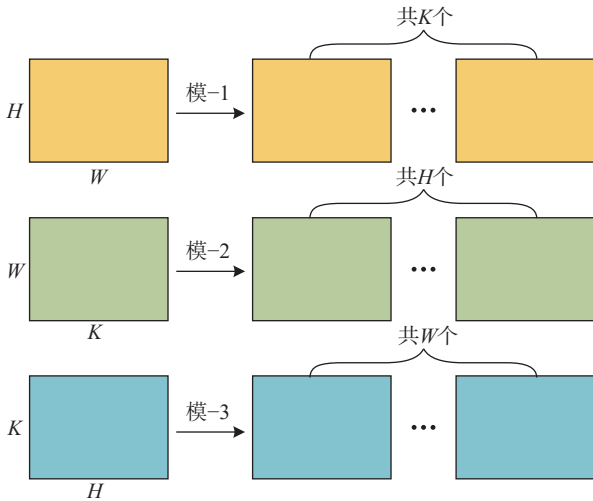


图 2 三阶张量展开示意图

### 1.2 张量乘法

张量乘法是指张量与矩阵做相乘运算。首先将张量展开为矩阵, 然后按照矩阵乘法运算法则与矩阵相乘, 最后再将相乘后的矩阵重新排列为张量。如图 2 所示, 由于三阶张量可按模-1、模-2 和模-3 展开, 因此, 按照不同模式展开后与矩阵相乘的结果并不相同。本文所涉及的数据维度最高为三维, 因此仍旧以三阶张量乘法为例。根据张量展开的模式, 张量乘法可表示为:

$$Y = X \times_n U \Leftrightarrow Y_{(n)} = UX \quad (1)$$

式中:  $n$  为张量展开的模式,  $Y_{(n)}$  为张量  $Y$  的模- $n$  展开。

### 1.3 张量的 CP 分解

张量的 CP 分解实质是将一个张量视为多个秩 1 张量线性表示的近似。以三阶张量  $X$  为例, CP 分解可表示为:

$$X \approx \sum_{r=1}^R \lambda_r a_r \circ b_r \circ c_r \quad (2)$$

式中:  $R$  是正整数,  $\lambda \in \mathbb{R}^R$  为维度  $R$  的矢量,  $a_r \in \mathbb{R}^W$ 、 $b_r \in \mathbb{R}^H$  和  $c_r \in \mathbb{R}^K$  为秩 1 张量, “ $\circ$ ” 为向量外积运算。三阶张量的 CP 分解示意图如图 3 所示。

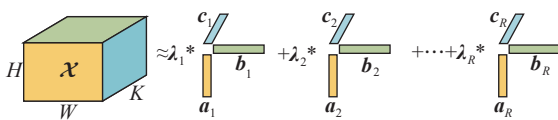


图 3 三阶张量 CP 分解示意图

近年来, 已有许多 CP 分解求解算法, 本文采用交替最小二乘法 (alternating least squares, ALS) 算法对三阶张量进行 CP 分解。对于三阶张量  $X$ , 给定参数  $R$ , 采用 ALS 算法进行 CP 分解的具体

流程如表 1 所示。表 1 中 “ $\circ$ ” 表示 Khatri-Rao 积运算, 矩阵  $A$ 、 $B$ 、 $C$  的列向量即为式 (2) 中的  $a_r$ 、 $b_r$  和  $c_r$ 。

表 1 CP 分解算法流程

算法1 ALS进行张量CP分解算法流程
1. 初始化矩阵 $A \in \mathbb{R}^{W \times R}$ 、 $B \in \mathbb{R}^{H \times R}$ 和 $C \in \mathbb{R}^{K \times R}$
2. 计算 $V = A^T A * B^T B * C^T C$
3. 计算 $A = X_{(1)}(C \circ B \circ A)V^\dagger$ , 归一化 $A$ , 求解 $\lambda$
4. 计算 $B = X_{(2)}(C \circ B \circ A)V^\dagger$ , 归一化 $B$ , 更新 $\lambda$
5. 计算 $C = X_{(3)}(C \circ B \circ A)V^\dagger$ , 归一化 $C$ , 更新 $\lambda$
6. 计算 $\varepsilon = \ X - A(\lambda \circ \lambda)(C \circ B)\ _F^2$ , 若 $\varepsilon$ 大于阈值, 则跳转到步骤 2; 若 $\varepsilon$ 小于阈值或达到最大迭代次数, 则运算终止

### 1.4 张量的 Tucker 分解

张量的 Tucker 分解可视为高维主成分分析 (principle component analysis, PCA), 能将张量分解为一个核心张量和相应的因子矩阵相乘的形式。以三阶张量为例, 张量  $X$  的 Tucker 分解可表示为:

$$X \approx \mathcal{G} \times_1 T^{(1)} \times_2 T^{(2)} \times_3 T^{(3)} \quad (3)$$

式中:  $\mathcal{G} \in \mathbb{R}^{O \times P \times Q}$  为核心张量,  $T^{(1)} \in \mathbb{R}^{W \times O}$ 、 $T^{(2)} \in \mathbb{R}^{H \times P}$  和  $T^{(3)} \in \mathbb{R}^{K \times Q}$  分别是与核心张量  $\mathcal{G}$  在 3 个维度上相乘的因子矩阵。三阶张量 Tucker 分解的示意图如图 4 所示。

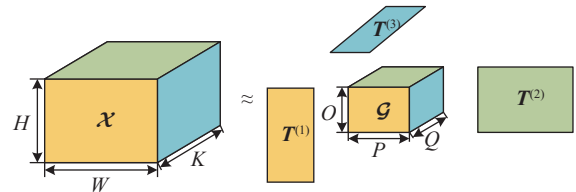


图 4 三阶张量 Tucker 分解示意图

张量的 Tucker 分解求解算法已有很多, 本文采用基于高阶正交迭代算法 (high-order orthogonal iteration, HOOI) 实现三阶张量的 Tucker 分解, HOOI 的算法流程如表 2 所示<sup>[12]</sup>。

表 2 HOOI 算法流程

算法2 HOOI算法流程
1. 初始化矩阵 $T^{(1)}$ 、 $T^{(2)}$ 和 $T^{(3)}$
2. 计算 $V = X \times_1 T^{(1)T} \times_2 T^{(2)T} \times_3 T^{(3)T}$
3. 对 $V_{(1)}$ 奇异值分解, 保留前 $O$ 个奇异向量以更新 $T^{(1)}$
4. 对 $V_{(2)}$ 奇异值分解, 保留前 $P$ 个奇异向量以更新 $T^{(2)}$
5. 对 $V_{(3)}$ 奇异值分解, 保留前 $Q$ 个奇异向量以更新 $T^{(3)}$
6. 计算 $\varepsilon = \ X - V\ _F^2$ , 若 $\varepsilon$ 大于阈值, 则跳转到 2; 若 $\varepsilon$ 小于阈值, 则运算终止
7. 迭代终止时, 令 $\mathcal{G} = V$

## 2 基于张量分解的红外监控视频压缩

### 2.1 基于张量的线性近似表示

基于张量分解的红外监控视频压缩的核心思想是利用张量分解算法对红外监控视频数据进行线性近似表示。假设将红外监控视频数据视为一个三阶张量 $\mathcal{X}$ ，则经张量分解后重建的张量 $\tilde{\mathcal{X}}$ 即为张量 $\mathcal{X}$ 的线性近似表示，可用下式表示：

$$\min_{\tilde{\mathcal{X}}} \|\mathcal{X} - \tilde{\mathcal{X}}\| \quad (4)$$

若采用 CP 分解，则张量 $\tilde{\mathcal{X}}$ 可表示为：

$$\tilde{\mathcal{X}} = \sum_{r=1}^R \lambda_r \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r \quad (5)$$

若采用 Tucker 分解，则张量 $\tilde{\mathcal{X}}$ 可表示为：

$$\tilde{\mathcal{X}} = \mathcal{G} \times_1 \mathbf{T}^{(1)} \times_2 \mathbf{T}^{(2)} \times_3 \mathbf{T}^{(3)} \quad (6)$$

### 2.2 参数选择与重建质量控制

对于张量 CP 分解，只需对式(2)中的参数  $R$  设置不同的值，即可调节重建质量与压缩后的数据量。一般而言，参数  $R$  越大，重建质量越高，压缩后数据体量也随之增大。

一段长度为 200 帧、尺寸为 512×640 的红外视频监控数据，采用张量 CP 分解，在不同参数  $R$  下的重建质量如图 5 所示。横轴代表参数  $R$ ，其取值为 10、30、50、70、90、110、130、150；纵轴为张量 CP 分解后的重建质量，用 PSNR 进行度量。从图 5 可以看出，张量 CP 分解后的重建质量随着参数  $R$  的增大而增大。而随着参数  $R$  的增大，表 1 中矩阵  $\mathbf{A}$ 、 $\mathbf{B}$ 、 $\mathbf{C}$  的尺寸随之增大，数据量也随之增大。因此，对于张量的 CP 分解，参数  $R$  的设置与比特消耗和重建质量呈正相关。

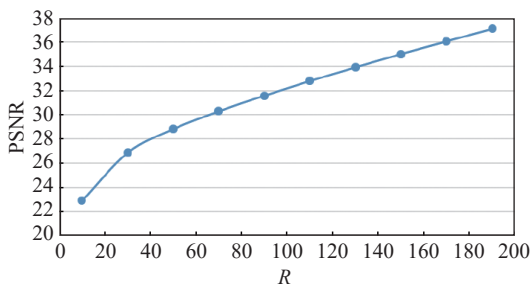


图 5 张量 CP 分解中参数  $R$  与重建质量变化趋势图

对于张量 Tucker 分解，需要分别调整式(3)中  $O$ 、 $P$ 、 $Q$  3 个参数，此 3 个参数分别控制式(3)中核心张量  $\mathcal{G}$  的 3 个维度及  $\mathbf{T}^{(1)}$ 、 $\mathbf{T}^{(2)}$ 、 $\mathbf{T}^{(3)}$  3 个因

子矩阵的维度，与压缩后的视频数据体量成正比。为了展示  $O$ 、 $P$ 、 $Q$  3 个参数对重建质量的影响，本文固定其中 1 个或 2 个参数，然后统计其他参数变化时重建质量的变化情况。在统计实验中，实验数据仍为长度 200 帧、尺寸 512×640 的红外视频监控数据。参数  $O$ 、 $P$ 、 $Q$  与重建质量变化情况如图 6 和图 7 所示。

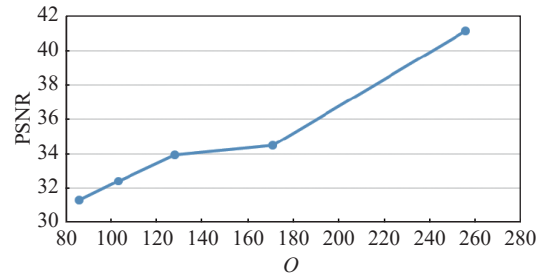


图 6 张量 Tucker 分解中参数  $O$  与重建质量变化趋势图

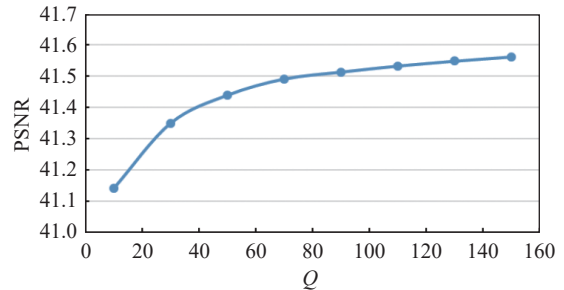


图 7 张量 Tucker 分解中参数  $Q$  与重建质量变化趋势图

在图 6 中，参数  $Q$  选定为 150，然后观测参数  $O$  和  $P$  同时取视频的高(512)和宽(640)的 1/2、1/3、1/4、1/5、1/6 时重建质量的变化情况。可以看出，在参数  $Q$  固定的情况下，同时改变参数  $O$  和  $P$  时，重建质量随着参数  $O$  和  $P$  增大而增大。

在图 7 中，参数  $O$  和  $P$  同时取视频高和宽的 1/2，参数  $Q$  取值为 10、30、50、70、90、110、130、150。可以看出，在参数  $O$  和  $P$  固定的情况下，重建质量随着参数  $Q$  增大而增大。与图 6 相比，图 7 中重建质量的增幅较小，参数  $Q$  在取值为 10 和 150 时，重建质量相差不超过 0.5 dB。尤其是在参数  $Q$  取值超过 100 时，重建质量的增幅小于 0.1 dB。因此，对于张量的 Tucker 分解，当固定其中 2 个参数时，另一个参数的变化对重建质量影响较弱。

通过以上分析可以得出：对于张量的 CP 分解，本文通过调节表 1 中参数  $R$  来控制重建质量；对于张量的 Tucker 分解，本文固定表 2 中的参数  $Q$ ，通过调整参数  $O$  和  $P$  实现重建质量的控制。

### 2.3 系数量化

基于张量分解的线性近似表示的求解,其结果需要用浮点数表示。如表1所示,需要对矩阵  $A$ 、 $B$ 、 $C$  归一化处理,进而引入浮点数的运算,然而浮点数的引入会降低压缩效率。为了提升压缩效率,本文采用标量量化器对张量分解得到的系数进行量化。同时,对张量 CP 分解在参数  $R$  取 90,张量 Tucker 分解  $O$ 、 $P$ 、 $Q$  3 个参数取 320、256、10 时得到的分解系数矩阵进行量化,进而观测不同量化步长下重建质量的变化趋势。

张量 CP 分解和 Tucker 分解的系数矩阵在不同量化步长(记为  $Q_{\text{step}}$ )下,重建质量的变化趋势。如图8和图9所示,横轴为量化步长  $Q_{\text{step}}$  的对数坐标,纵轴为重建质量。

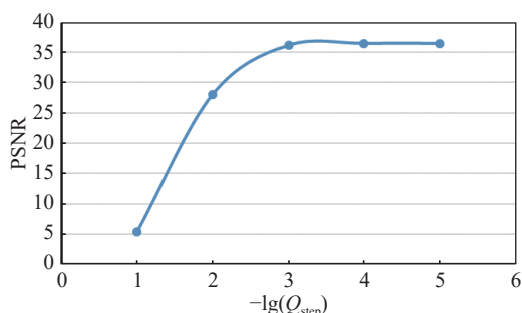


图8 张量 CP 分解系数矩阵量化步长对重建质量影响图

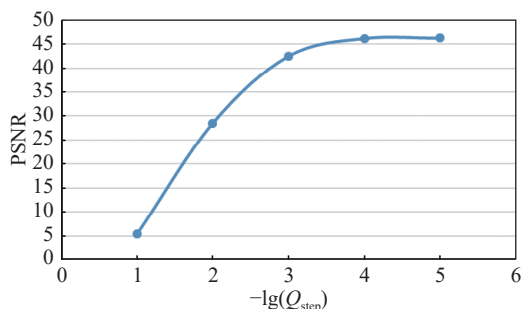


图9 张量 Tucker 分解系数矩阵量化步长对重建质量影响图

从图8可以看出,当量化步长不超过 0.001 时,张量 CP 分解的重建质量变化较为平缓,维持在 36.1 dB 以上;当量化步长超过 0.001 时,重建质量大幅下降,在量化步长为 0.1 时,重建质量仅为 5.3 dB。

从图9可以看出,当量化步长不超过 0.0001 时,张量 Tucker 分解的重建质量变化较为平缓,维持在 46.2 dB 以上;当量化步长为 0.001 时,重建质量为 42.5 dB;当量化步长超过 0.01 时,重建

质量大幅下降,在量化步长为 0.1 时,重建质量仅为 5.3 dB。

从以上分析可以得出:对于张量的 CP 分解,量化步长可选 0.001,可用 10 bit 描述系数矩阵中的元素;对于张量的 Tucker 分解,量化步长可选 0.0001,可用 14 bit 描述系数矩阵中的元素,对于核心张量中的系数则用 16 bit 进行描述。

### 2.4 比特消耗度量

根据上述系数量化及张量 CP 分解和 Tucker 分解系数矩阵的尺寸,根据模型参数即可计算比特消耗。对于张量 CP 分解,需要存储表1中的矩阵  $A$ 、 $B$ 、 $C$  及  $\lambda$ 。由于每个矩阵中每个元素经量化后可用 10 bit 表示,因此张量 CP 分解的比特消耗可通过下式计算得到:

$$rate_{CP} = \frac{(W+H+K+1) \times R \times 10}{W \times H \times K \times 8} \quad (7)$$

式中:  $rate$  为码率,表示单位像素的比特消耗。

对于张量的 Tucker 分解,需要存储表2中的核心张量  $\mathcal{G}$  和 3 个系数矩阵  $T^{(1)}$ 、 $T^{(2)}$  和  $T^{(3)}$ ,其比特消耗可通过下式计算得到:

$$rate_{Tucker} = \frac{O \times P \times Q \times 16 + (W \times O + H \times P + K \times Q) \times 15}{W \times H \times K \times 8} \quad (8)$$

## 3 实验结果

### 3.1 实验设置

实验数据来源于采集的红外机场监控视频,尺寸为  $512 \times 640$ ,帧率为 30,比特深度为 24。从中随机选取 4 段 200 帧的视频序列(分别记为  $I_1$ 、 $I_2$ 、 $I_3$ 、 $I_4$ )作为实验数据。4 段视频的第一帧如图10所示,可以看出,  $I_1$  内容较简单;  $I_2$  图像内容更复杂,但运动较少;  $I_3$  内容为飞机滑行,运动较大;  $I_4$  内容最复杂。需要指出的是,在测试过程中,本文首先将每段视频数据转化为 8 bit 的视频数据,然后对 8 bit 的红外视频进行压缩并观测各算法的压缩性能。

本文采用了张量的 CP 分解和 Tucker 分解两种算法。在测试过程中,张量 CP 分解中的参数  $R$  分别设为 40、55、80、115,量化步长  $Q_{\text{step}}$  为 0.001;张量 Tucker 分解中的 3 个参数  $O$  分别设为 320、213、160、106,  $P$  分别设为 300、277、220、177,  $Q$  分别设为 2、2、2、2,量化步长  $Q_{\text{step}}$  为 0.0001。

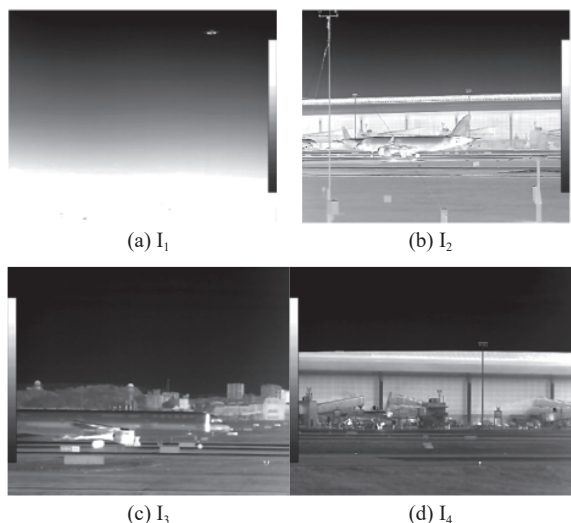


图 10 4 段测试视频第一帧截图

为了验证算法的有效性，本文以 HEVC 测试模型 HM-16.22 为基准。另外，考虑到张量表述需要预先缓存视频数据，与视频编码中随机接入 (random access, RA) 的应用场景类似，因此采用 RA 编码结构，编码参数采用该配置下的默认设置<sup>[16]</sup>。每段视频数据采用 4 个量化参数 (quantization parameter, QP) 进行压缩，分别为 22、27、32、37。

### 3.2 性能对比

为了全面客观地呈现所提算法的压缩性能，本文采用 BD-rate 和 BD-PSNR 作为评价指标。BD-rate 表示在同等客观质量下，测试方法相对于基准算法的码率节省，用百分比表示。BD-rate 取正值时表示率失真性能损失，取负值时表示率失真性能改善。在本文中，视频重建的客观质量用 PSNR 度量。BD-PSNR 表示在同等码率下，测试方法相对于基准算法的客观质量提升。

4 段测试视频序列的率失真曲线对比结果如图 11 所示。从图 11(a)中可以看出，在低码率下，张量 CP 分解和 Tucker 分解的压缩性能优于 HM-16.22。在同等质量下，基于张量分解的压缩方法其比特消耗更少。尤其是张量的 Tucker 分解，相比于 HM-16.22 可节省 77.7% 的 BD-rate。但对于测试视频 I<sub>2</sub>，基于 CP 分解的压缩方法性能较低，基于 Tucker 分解的编码性能优于 HM-16.22。对于测试视频 I<sub>3</sub>，CP 分解的性能依旧不理想，Tucker 分解的编码性能在低码率下优于 HM-16.22，当重建质量接近 38 dB 时，Tucker 分解所需的比特消耗约为 HM-16.22 的 1/100。对于测试视频 I<sub>4</sub>，CP 分解和 Tucker 分解的编码性能较差，

在高码率情况下，BD-PSNR 降低超 10 dB，原因在于测试视频 I<sub>4</sub> 的内容复杂，运动较多。

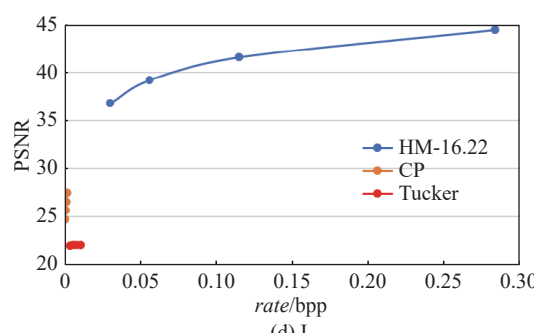
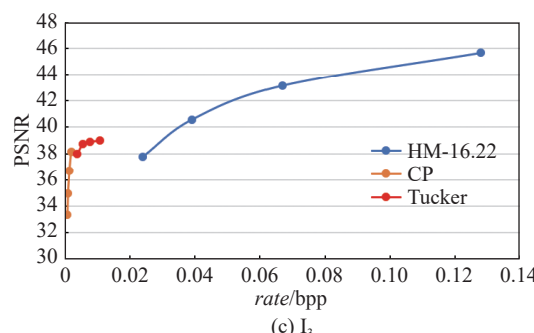
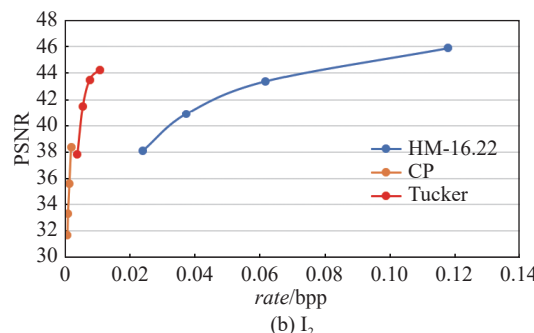
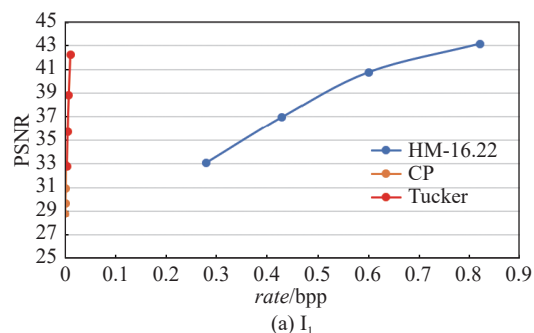


图 11 率失真曲线对比图

从以上结果可得出以下 3 点：基于张量分解的压缩算法适用于低码率的情况；Tucker 分解的压缩性能优于 CP 分解；张量分解适用于内容相对简单、存在大量平缓区域的视频。其原因在于基于表 1 的 CP 分解和基于表 2 的 Tucker 分解更适用于低秩+稀疏的数据类型。而相对简单、存在大量平缓区域的视频更符合低秩+稀疏的数据特点，

实验结果充分说明了这一点。

虽然采用基于张量分解的压缩性能仅适用于内容简单的视频,应用范围较小,但其压缩时间短。本文测试服务器的CPU为Intel Xeon R,内存为128 GB。在MATLAB 2016b上运行表1和表2实现张量的CP分解和Tucker分解。张量CP分解、Tucker分解和HEVC的编码时间如表3所示。可以看出,张量的CP分解的平均编码时间仅为HM-16.22的0.57%,Tucker分解的平均编码时间仅为HM-16.22的2.25%。对于解码时间,Tucker分解可实现实时解码,而CP分解的解码时间较长,平均需要60 s完成300帧视频的解码。

表3 编码时间对比

测试序列	类别		
	CP分解	Tucker分解	HM-16.22
I <sub>1</sub>	10.18	49.05	1732.13
I <sub>2</sub>	9.93	30.75	1823.10
I <sub>3</sub>	9.54	51.28	1830.22
I <sub>4</sub>	10.27	55.64	1897.19
平均时间/s	9.98	46.68	1820.70
时间比率/%	0.55	2.56	100.00

#### 4 结束语

本文提出了一种基于张量分解的红外机场监控视频压缩方法,首先将红外视频数据视为一个三阶张量,然后采用张量分解近似表示,得到的系数矩阵量化后采用HEVC帧内无损压缩方式进行压缩。实验结果表明,基于张量分解的算法适用于低码率、视频内容存在大量平缓区域的视频。另外,Tucker分解的压缩性能优于CP分解;对于内容简单的视频,相较于HEVC,Tucker分解可节省77.7%的BD-rate,编码性能大幅提升。基于张量分解的压缩算法的编码时间大幅减低,CP分解的压缩时间仅为HEVC的0.55%,Tucker分解的编码时间为HEVC的2.56%。

#### 参考文献

[1] WIEGAND T, SULLIVAN G J, BJONTEGAARD G, et al. Overview of the H. 264/AVC video coding standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 560–576.

[2] SULLIVAN G J, OHM J R, HAN W J, et al. Overview of the high efficiency video coding (HEVC) standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(12): 1649–1668.

[3] BROSS B, CHEN J, OHM J R, et al. Developments in international video coding standardization after AVC,

with an overview of versatile video coding (VVC)[J]. Proceedings of the IEEE, 2021, 109(1): 1–31.

[4] LI T, XU M, TANG R, et al. DeepQTMT: A deep learning approach for fast QTMT-based CU partition of intra-mode VVC[J]. IEEE Transactions on Image Processing, 2021, 30: 5377–5390.

[5] CHO S, KIM M. Fast CU splitting and pruning for suboptimal CU partitioning in HEVC intra coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23(9): 1555–1564.

[6] PAUL S, NORKIN A, BOVIK A C. Speeding up VP9 intra encoder with hierarchical deep learning-based partition prediction[J]. IEEE Transactions on Image Processing, 2020, 29: 8134–8148.

[7] KIM J, BLASI S, DIAS A S, et al. Fast inter-prediction based on decision trees for AV1 encoding[C]//IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Brighton: IEEE, 2019: 1627–1631.

[8] YUAN M, XUE Y, OHN S, et al. Fast CU size and PU partition decision for AVS2 intra coding[C]//IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB). Valencia: IEEE, 2018: 1–5.

[9] MANTEL C, FORCHHAMMER S. Compression of infrared images[C]//IS & T Symposium on Electronic Imaging: Science and Technology. California: IEEE, 2017: 21–26.

[10] CHAN C H, TENG C, CHUANG J H. Multi-intensity illuminated infrared video compression using MV-HEVC and 3D-HEVC[C]//IEEE International Conference on Multimedia and Expo Workshops (ICMEW). Hong Kong: IEEE, 2017: 10–14.

[11] BELYAEV E, FORCHHAMMER S. Drone HDR infrared video coding via aerial map prediction[C]//IEEE International Conference on Image Processing. Athens: IEEE, 2018: 1733–1736.

[12] DAS S. Hyperspectral image, video compression using sparse tucker tensor decomposition[J]. IET Image Processing, 2021, 15(4): 964–973.

[13] AIDINI A, TSAGKATAKIS G, TSAKALIDES P. Tensor decomposition learning for compression of multidimensional signals[J]. IEEE Journal of Selected Topics in Signal Processing, 2021, 15(3): 476–490.

[14] MATVEEV S, KURILOVICH A. Utilization of tensor decomposition for video-compression[C]//33rd International Conference on Computer Graphics and Vision. Moscow: Institute of Control Sciences of Russian Academy of Sciences, 2023, 582–589.

[15] KOLDA T G, BADER B W. Tensor decompositions and applications[J]. SIAM Review, 2009, 51(3): 455–500.

[16] BOSSEN F. Common test conditions and software reference configurations [R]. Geneva: Joint Collaborative Team on Video Coding (JCT-VC), 2013.