

面向 6G 的全景视频片划分优化编码算法

杨桃雨, 徐媛媛, 谭增洁

引用本文

杨桃雨, 徐媛媛, 谭增洁. 面向 6G 的全景视频片划分优化编码算法[J]. 计算机科学, 2022, 49(6): 66-72.

YANG Tao-yu, XU Yuan-yuan, TAN Zeng-jie. [Tile Partition Optimized Omnidirectional Video Coding for 6G Network](#)[J]. Computer Science, 2022, 49(6): 66-72.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[视频隐私保护技术综述](#)

Review on Video Privacy Protection

计算机科学, 2022, 49(1): 306-313. <https://doi.org/10.11896/jsjcx.201200047>

[基于深度强化学习的无人机辅助弹性视频多播机制](#)

Deep Reinforcement Learning Based UAV Assisted SVC Video Multicast

计算机科学, 2021, 48(9): 271-277. <https://doi.org/10.11896/jsjcx.201000078>

[异构无线网络中基于非正交多址的可伸缩视频多播机制](#)

Non-orthogonal Multiple Access Enabled Scalable Video Multicast in HetNets

计算机科学, 2021, 48(11): 356-362. <https://doi.org/10.11896/jsjcx.200900080>

[基于运动矢量细化的帧率上变换与 HEVC 结合的视频压缩算法](#)

Video Compression Algorithm Combining Frame Rate Up-conversion with HEVC Standard Based on Motion Vector Refinement

计算机科学, 2020, 47(2): 76-82. <https://doi.org/10.11896/jsjcx.190500092>

[基于深度残差网络的 HEVC 压缩视频增强](#)

Deep Residual Network Based HEVC Compressed Videos Enhancement

计算机科学, 2019, 46(3): 88-91. <https://doi.org/10.11896/j.issn.1002-137X.2019.03.011>

面向 6G 的全景视频片划分优化编码算法

杨桃雨 徐媛媛 谭增洁

河海大学计算机与信息学院 南京 211100

(yty@hhu.edu.cn)

摘要 第 6 代无线通信的兴起为虚拟现实全景视频的发展提供了更广阔的前景。基于片的全景视频编码方案能够在相同网络带宽条件下提升全景视频的观看体验,片划分大小会影响视频传输的性能,与较小的片相比,采用较大的片能够降低编码冗余,但需传送更大区域来覆盖相同视口,这会造成更多的像素开销。目前已有片划分工作主要是针对矩形视口而设计的,而实际中,球型视频到二维平面的投影会对视口不同区域进行不同程度的拉伸,用矩形片覆盖不规则视口区域时像素开销计算更为复杂。针对这一挑战,文中提出了一种适用于用户真实视口的片划分算法。首先,分析了投影格式对视口区域造成的拉伸失真,并对在不同片划分下不规则视口区域的额外像素开销进行了推导;然后,通过权衡不同切片粒度编码单元的像素开销与编码效率的关系,提出了一种全景视频序列最优片划分编码方案;最后,通过与片划分穷举搜索法进行对比,验证了所提算法在较少的计算复杂度下能取得与穷举搜索法相当的传输效率。

关键词: 6G;感知互联;全景视频;视频编码;片划分

中图法分类号 TP751.1

Tile Partition Optimized Omnidirectional Video Coding for 6G Network

YANG Tao-yu, XU Yuan-yuan and TAN Zeng-jie

School of Computer and Information, Hohai University, Nanjing 211100, China

Abstract The rise of the 6G wireless communication provides a broader prospect for the development of virtual reality panoramic video. The tile-based panoramic video coding scheme can improve the viewing experience of 360° video under the same condition of network bandwidth. Tile partition affects the video transmission performance. Compared with small tiles, using large tiles can effectively improve coding efficiency, but it will cause more pixel overhead by transmitting a larger area to cover the viewport. The existing tile partition work is mainly designed for the rectangular viewport, but the projection of spherical video to a two-dimensional plane will stretch different areas of the viewport to varying degrees in practice. Deriving the pixel overhead associated with covering irregular viewport areas with rectangular tiles is more complicated. To address this challenge, a tile partitioning algorithm for the user's real viewport has been proposed in this paper. Firstly, the stretching distortion of the viewport caused by the projection format is analyzed, and the pixel overhead of the irregular viewport with different tile partition sizes is derived. Secondly, by trading off the pixel overhead and the coding efficiency of tiles with different granularity, an optimal tile partition scheme has been proposed for the panoramic video sequence. Finally, the proposed scheme is compared with the exhaustive search method for tile partition in the experiment, and the results show that the proposed algorithm can achieve almost the same transmission efficiency as the exhaustive search method with less computational complexity.

Keywords 6G networks, Sensory interconnection, Omnidirectional video, Video coding, Tile partition

1 引言

未来数字城市万物互联,随着智能终端和数字媒体用户数量的增加,多媒体的表现形式不断丰富,从最初黑白视频到高清视频再到全景视频,多样、智能的数字视频形式给人们带来了沉浸且愉快的观看体验。5G 商用标志着移动通信真正迈入互联时代,而当前在全球兴起的第六代(6G)无线通信

网络系统使得更多应用成为可能,如原生 AI、感知互联技术等^[1-3]。其中,增强型移动宽带(Enhanced Mobile Broadband, eMBB+)对峰值速率、用户体验速率、端到端时延等提出了更高的要求^[4],为虚拟现实、增强现实等沉浸式体验应用提供了更好的支持。

全景视频,作为虚拟现实的一种实现方式,允许用户随意选择观看角度,为用户提供身临其境的观看体验。实现虚拟

到稿日期:2022-04-02 返修日期:2022-04-24

基金项目:国家自然科学基金(61801125)

This work was supported by the National Natural Science Foundation of China (61801125).

通信作者:徐媛媛(yuanyuan_xu@hhu.edu.cn)

现实沉浸式体验需要接近人类感知极限的高分辨率、高帧率的视频予以支撑^[5-6],用户观看全景视频时,可透过头戴式显示设备(Head Mounted Display, HMD)的放大镜片观看有限视口区域的内容^[7-8],由于视口区域内容被放大观看,因此全景视频的分辨率和帧率远高于传统视频。此外,为了避免头部快速旋转引起的晕眩感,对运动到光子(Motion-to-Photon, MTP)的延迟亦有严格的延迟约束^[9-11],因此全景视频需要低延迟的高比特率传输。针对这一挑战,需要为全景视频设计一种高效的压缩编码方案,以便在相同传输速率下实现更好的视频体验。

由于用户在同一时刻只能看到有限视口(Field of View, FoV)区域的内容,因此使用 FoV 位置信息可以有效提高全景视频的传输效率。高效视频编码(HEVC)标准的编码工具可以将图像编码成多个可独立解码的片(tile),通过仅传输与 FoV 相关的片来降低对带宽的要求^[12]。与传统的平面视频相比,全景视频呈现于球面上。由于缺乏球型视频的编码方法,因此现有的全景视频编码通常先将球面内容投影到二维平面上,再采用传统的视频编码方案进行编码。现有的一些全景视频编码的工作在码率分配^[13-15]、失真度量^[16-19]、质量评估^[20]方面考虑了投影格式的影响,对全景视频的编码参数进行了优化。

除了片的编码参数优化,片划分方案亦会对全景视频的传输效率产生影响。采用较大的片能够提高片的编码效率,但需传送更大区域来覆盖相同视口,这会造成更多的像素开销,如何选择划片大小以提高传输效率是值得研究的问题。目前已有片划分工作^[21-28]主要是针对矩形视口而设计的,而实际的视口中球型视频到二维平面的投影形状并不规则。RAPT360^[28]设计了不规则视口的非均匀片划分方案,但该方案只在3种片划分颗粒度中进行选择,应尽量用细颗粒度的片来覆盖边界区域。

本文针对视口区域不规则的特点,为片编码参数可依据视口信息及视口预测误差进行优化的全景视频编码方案^[15]设计了最佳的均匀片划分方案,从而提高了视口区域的传输效率。本文的主要贡献如下:

(1)对常见的等矩形投影(ERP)格式下二维平面上的不规则视口区域进行了表示,并对不同片划分下传送视口区域造成的额外像素开销进行了推导。

(2)对编码参数优化的全景编码方案^[15]在相同加权失真下的划片数目和比特率之间的关系进行了曲线拟合。

(3)通过权衡额外像素开销和编码效率,提出了一种最优均匀片划分方案来提高视口区域的传输效率。通过与穷举搜索法进行对比实验,验证了所提片划分方案的有效性。

2 相关工作

在全景视频的编码参数优化方面,现有的一些全景视频编码工作考虑了投影的影响。由于全景视频投影后的失真会受到投影方式的影响,例如在 ERP 格式下球面视频在赤道附近拉伸较小,两极处拉伸更大,因此普通视频失真度量的方式无法反映投影后全景视频的失真。为了更好地提升编码效率,考虑到不同投影格式的平面图像像素在球域中具有不同

的影响,Li 等^[13]提出了一种有效比特率分配的速率控制方法,该方法对码率分配中全景视频的编码参数进行了优化。Zhou 等^[14]通过观察,发现编码树单元(Coding Tree Unit, CTU)的熵或比特率与其几何位置相关,因此提出了一种新的熵平衡优化方法来增强全景视频的编码性能。通过计算 CTU 的熵和比特率及其在球形空间中几何位置的相关参数,可有效增强全景视频的编码性能。Li 等^[16]、Liu 等^[17]及 Guo 等^[29]针对球域投影转换引入的像素失真,通过率失真优化计算重建后像素的加权失真,以得到的权重更新拉格朗日乘子(λ)。Chen 等^[30]考虑增强现实中用户观看信息的重要性,对非用户关注的视频内容进行低码率压缩,从而达到节约能耗并保证用户观感的目的。文献^[20]提出了一种基于球形加权的质量评估指标来度量失真,该指标根据不同投影格式计算投影后二维投影平面上不同区域的权重,以此减小投影格式对编码造成的影响,该方式能够较好地衡量全景视频编码造成的失真。文献^[15]考虑了视口预测误差,按照 tile 出现在视口的概率对 tile 的编码参数进行了优化,在相同比特率下提升了视口区域的期望重建质量。上述工作以降低全景视频投影失真为优化目标,提升了用户观看区域内的视觉质量,但是这些工作未考虑视频编码过程中编码单元尺寸对传输效率的影响。

在全景视频的片划分方面,大多数工作为矩形的视口区域设计了最佳的片划分方案。由于固定的片划分机制不能提供适合于用户视口的片粒度,发送内容中包含了用户视口以外的区域,因此导致了冗余数据。较小的 tile 可以为用户视口提供更细小的粒度以减少冗余像素,但相应地也会降低编码效率^[31]。Guilherme 等^[21]提出了基于机器学习的显著性检测方法,并设计了一种考虑视频显著性和空间信息的自适应编码方案,对矩形视口不同片划分下的额外像素开销进行了推导,并以此得出最优的片划分粒度,通过预测的视场信息对与视口有关的片进行优先传输,能够有效降低带宽消耗,但是该方案对预测视口精度有较高的要求。Xiao 等^[22]提出了一种平衡视频存储成本和带宽成本的片划分方案,从节省带宽的角度考虑了全景视频中不可见区域传输造成的带宽浪费,根据每个片的存储成本,应用整数线性规划模型来求解最优片选择问题,能够为不同的视频选择不同的片划分大小。该方案选择了4种固定片粒度,并将其作为可选项为视口区域来选择适合不同视频序列的片尺寸,但并未考虑编码中片粒度的影响,仅以降低传输成本为目的。文献^[23]通过分析区域的重要性,基于 H.264/AVC 编码框架自适应地选择片的尺寸及版本,以调整采样密度来确定 tile 的大小,能够在固定码率下提升全景视频的编码质量。PANO 方案^[25]根据可察觉失真为矩形视口设计了片划分方案,为 360° 视频建立新的质量模型,提出了一种可变尺寸的平铺方案,以达到感知质量和视频编码效率之间的平衡。Sanchez 等^[26]根据时空特性,对不同 tile 采用统一且固定编码参数的编码方案,然后对划片数目和节约比特率之间的关系进行了拟合,以期提升编码效率。

在片尺寸选择中,文献^[28,32]均选择了3种不同大小的片划分方案,为不规则视口设计了非均匀片划分方案,选择时

尽量用细粒度的片来覆盖边界区域。相比非均匀划片,均匀划片更为常用。文献[33]则提出了一种基于计算几何方法的自适应片划分方案 VAD360,以视频帧上的视觉注意信息为输入,在帧上提供合适的非重叠可变尺寸 tile 来覆盖视口。文献[34]在片尺寸选择时使用了穷举的方法,该方法基于视频流传输的场景,解决了网络需求增加导致全景视频内容在空间访问时复杂度升高的问题,设计了对于传输来说最佳片的尺寸选择方案。

现有的片划分方案大多从流传输角度出发,为规则的矩形视口区域选择适宜的片尺寸,以提高传输效率,而实际中,受到投影格式影响,真实视口呈不规则形状,因此需要针对不规则视口区域来设计最优片划分方案以减少比特浪费。此外,片划分的编码工作仅以 3~4 个片粒度来权衡传输速率与片粒度之间的关系,所得结果并不能代表最优片划分结果。上述工作均未考虑如何为不规则视口设计均匀划片方案。

3 片划分最优的全景视频编码方案

3.1 不规则视口的像素开销

由于缺乏球型视频的编码方法,现有的全景视频编码需要先将球面内容投影到二维平面再进行编码,其中 ERP 是最为常用的投影方式。球面上近似矩形的视口经 ERP 投影格式投影,不同纬度的区域拉伸程度不同将产生不规则区域,如图 1 所示。

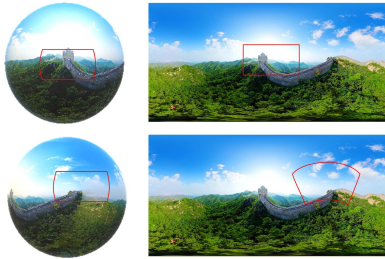


图 1 视口投影变换示例

Fig. 1 Example of viewport projection

假如视口在半径为 R 的球面上可由极坐标分别为 (ϕ_1, θ) , (ϕ_2, θ) , (ϕ_{11}, θ_1) 和 (ϕ_{22}, θ_1) 的 4 个点张开的矩形区域进行表示,其中 θ 和 θ_1 分别为视口端点靠近和远离极点的极角, $\phi_1, \phi_2, \phi_{11}$ 和 ϕ_{22} 分别为视口 4 个端点的方位角。上述变量满足式(1)和式(2)。

$$\phi_1 + \phi_2 = \phi_{11} + \phi_{22} \quad (1)$$

$$(\phi_1 - \phi_2)R \cos \theta = (\phi_{11} - \phi_{22})R \cos \theta_1 = L \quad (2)$$

下文对视口在球面上的边界进行表示。假如用 A 和 B 表示视口靠近极点的两个端点 (ϕ_1, θ) 和 (ϕ_2, θ) ,那么 A 点和 B 点在球面上的最短连线(即视口的边界)并不在纬度为 θ 的纬线上,而是落在以球心 O 点为圆心、 A 点和 B 点确定的圆周上。如图 2 所示,假如 A 点和 B 点在球面上的最短连线中点用 D 点表示, A 点和 B 点在球体内部的中点用 C 点进行表示,那么 C 点的直角坐标为

$$\left(\frac{R \cos \theta (\cos \phi_1 + \cos \phi_2)}{2}, \frac{R \cos \theta (\sin \phi_1 + \sin \phi_2)}{2}, R \sin \theta \right)。$$

通过 O 点和 C 点连线的延长线穿透球面可以得到 D 点的

坐标。位于 O 点和 C 点连线上且落在球面上的 D 点直角坐标满足如下关系:

$$a^2 (R^2 \cos^2 \theta \left(\frac{\cos \phi_1 + \cos \phi_2}{2} \right)^2 + R^2 \cos^2 \theta \left(\frac{\sin \phi_1 + \sin \phi_2}{2} \right)^2 + R^2 \sin^2 \theta = 1 \quad (3)$$

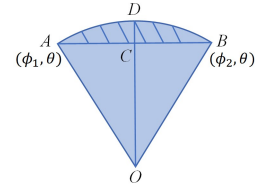


图 2 视口在球面上的边界

Fig. 2 Viewport boundary on the sphere

由式(3)可以得到常数 a 的值,那么靠近极点视口水平方向边界中点 D 的极坐标 (ϕ', θ') 如下:

$$\theta' = \arcsin \theta \left(\frac{1}{a} R \sin \theta \right) \quad (4)$$

$$\phi' = \frac{\phi_1 + \phi_2}{2} \quad (5)$$

按照上述方法,可以类似地计算出上下边界上每个点的极坐标。

经过 ERP 投影后,整个球形画面被投影到像素分辨率为 $m \times n$ 的二维平面上,其中 m 和 n 分别是图像每行、每列的像素数目。以纬度为 θ 的纬线上长度为 L 的线段为例,可以得到该纬线片断在二维平面上水平方向占据的像素数目 W_1 。

$$W_1 = \frac{mL}{2\pi R \cos \theta} \quad (6)$$

由式(6)可以看出,长度相同的纬线片断占据的像素数目与纬度相关。视口远离极点的两个端点确定的纬线在水平方向占据的像素数目 W_2 为:

$$W_2 = \frac{mL}{2\pi R \cos \theta_1} \quad (7)$$

长度为 L' 的经线投影后在垂直方向占据的像素数据为:

$$H = \frac{nL'}{\pi} \quad (8)$$

由式(8)可以看出,长度相同的经线片段占据的像素数目与经度无关。因此,ERP 投影格式仅对纬度不同的视口区域做了不同程度的拉伸。二维平面上不规则视口示意图如图 3 所示。

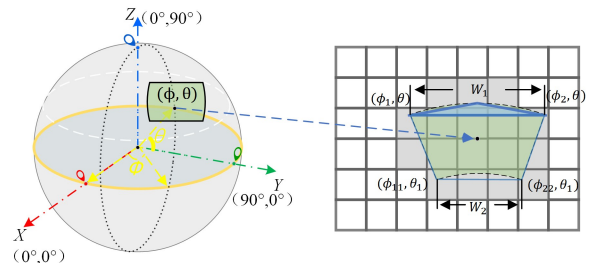


图 3 二维平面上不规则视口示意图(电子版为彩图)

Fig. 3 Illustration of irregular viewport on 2D plane

通过球面视口边界的表示及 ERP 投影变换的分析,下文将对不规则视口像素开销进行推导。

假如片大小为 $t \times s$,其中 t 和 s 分别是片每行、每列的

像素数目。文献[9]对不同片划分下一维以及二维矩形区域的像素开销进行了推导。在一维情形下,如果用 t 和 R 分别表示片和区域的长度,那么期望传送的像素数目 P 可以表示成:

$$E\{P\}=R+t-1 \quad (9)$$

那么,根据文献[9],在水平方向上长度为 W_1 和 W_2 的边界需要传送的期望像素个数分别为:

$$E\{P_{w1}\}=W_1+t-1 \quad (10)$$

$$E\{P_{w2}\}=W_2+t-1 \quad (11)$$

垂直方向上长度为 H 的边需要传送的期望像素个数为:

$$E\{P_h\}=H+s-1 \quad (12)$$

不规则的二维视口的像素开销的计算过程如下。如图3所示,真实视口在二维平面上的视口形状不规则,为了降低计算复杂度,我们忽略了靠近赤道边界产生的不太明显的形变,并将靠近极点的形变区域近似用三角形表示(即图3中的蓝色斜线区域)。该区域在垂直方向需要传送的 tile 的数量 N' 由边界中点和边界的高度差 $|\theta' - \theta|$ 决定。

$$N' = \begin{cases} 0, & H \leq P_h + t - 1 \\ \lceil \frac{|\theta' - \theta|}{t} \rceil, & H > P_h + t - 1 \end{cases} \quad (13)$$

而对应该区域编码 tile 包含的像素个数的期望 $E\{N'\}$ 为:

$$E\{N'\} = t \times \left(\frac{E\{P_{w1}\} \times N'}{2} \right) \quad (14)$$

最终,可得一个平面对 FoV 编码的 tile 的面积。

$$E\{P\} = \frac{1}{2} \{ (E\{P_{w1}\} + E\{P_{w2}\}) \times E\{P_h\} \} + E\{N'\} \quad (15)$$

本文将投影后不规则视口扇形区域的精确面积记为 S_f , 可通过前文球面上的视口表示、ERP 格式转换公式求积分得到。根据 FoV 区域总传输面积与 FoV 精确面积的比值,不规则 FoV 区域的像素额外开销为:

$$\phi_{(t,s)} = \frac{E\{P\}}{S_f} \quad (16)$$

额外传输开销 $\phi_{(t,s)}$ 为大于或等于 1 的数值,数值越大,表示该片划分下传输了更多的视口外像素。

3.2 不同片划分下的编码效率

文献[14]对划片数目和节约比特率之间的关系进行了拟合,但该工作只针对所有片采用统一且固定编码参数的情况,不适用于片编码参数可依据视口及视口预测差信息进行优化的编码方案[15]。该方案按照 tile 出现在视口中的概率对片的量化参数进行优化,而不是采用固定、统一的编码参数,划片数目和比特率间的拟合过程更为复杂。本文对此情况下划片数目和比特率间的关系进行了拟合。为了降低复杂度,我们只考虑片为正方形($s=t$)的情况。

首先,我们选取了多个不同片划分大小的划分方案,在每种划分大小下对视频序列分别采用基础量化步长{22,27,32,37}进行编码,每个片的编码参数根据编码方案[15]进行计算。

采用不同片大小的编码方案时整个视频画面的质量不同,我们画出每种片大小下的球形加权失真和比特率曲线。在球形加权率失真曲线上,选择了相同的平均球形加权失真,得到了对应的比特率。根据划片数目 $T(T=mn/s^2)$ 和得到的比特率 $\eta(T)$ 采用如下方式对它们的关系进行拟合。

$$\eta(T) = b \ln(T) + \alpha \quad (17)$$

其中, α 表示比特率-划片数目曲线的修正值, α 和 b 都是与序列相关的常数。

3.3 片划分最优的编码方案

通过权衡额外像素传输开销与编码效率之间的关系,我们最终可以按照式(18)得到传输性能最优的划片。

$$\min_s \eta(mn/s^2) \times \phi(s,s) \quad (18)$$

对于任意的全景视频序列,首先根据视点位置计算出不规则 FoV 区域的像素开销 $\phi(s,s)$; 然后通过采用若干个有限的片划分方案对序列编码,得到序列编码效率和划片大小的拟合参数,从而得到 $\eta(mn/s^2)$; 最后将所有可能的划片方案都放入式(18),得到传输像素开销与比特率乘积最小的划片方案,即最佳的划片方案。

得到最佳的划片大小后,根据文献[15]提出的量化步长调整方案计算每个片的量化步长,用计算出的量化步长对每个片进行编码,这样就完成了该全景视频的片划分最优的编码过程。

4 实验

4.1 实验设置

实验使用 HEVC 标准参考软件 HM16.16 实现编码方案,实验序列为 JVET 提供的分辨率为 3840×1920 的测试视频序列,即 Kiteflite, Harbor, Snowberg 以及 Trolley, 编码配置为 Random Access 方式。如表 1 所列,在实验中共考虑了 7 种划片方案,并使用片划分穷举搜索算法进行对比实验来验证算法的有效性。

表 1 不同片划分方案的片尺寸

Table 1 Tile size for different tile partition schemes

划片总数	尺寸(度数)	尺寸(像素)
8	90°×90°	960×960
18	60°×60°	640×640
32	45°×45°	480×480
72	30°×30°	320×320
98	25°×25°	224×224
162	20°×20°	213×213
200	18°×18°	192×192

4.2 实验结果分析

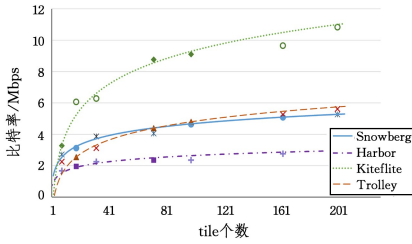
首先验证了适用于编码方案[15]的比特率和编码参数拟合曲线的拟合效果。我们选取了划片总数为 18,72,98 这 3 种划分方案,在同一加权失真下(Kiteflite, Trolley 序列取 WS-PSNR=32 dB,其余序列取 WS-PSNR=37 dB)获得每种划片数目对应的比特率,并依据式(17)进行拟合,可得到拟合曲线相关参数值。表 2 列出了拟合后 4 个全景视频序列的 R-T 曲线中 b 和 α 的数值,比特率单位为 Mbps。图 4 给出了比特率关于划片数目的拟合曲线,图中每条曲线上仅有 3 个实心点表示拟合时所使用的实验数据。

表 2 比特率关于划片数目拟合曲线的参数

Table 2 Parameters in fitted curve of bit rate on number of tiles

	Kiteflite	Harbor	Snowberg	Trolley
b	0.85	0.39	2.39	1.20
α	0.77	0.88	-1.25	-0.83

然后使用其余划片方案,即划片总数分别为 8, 32, 162, 200 的方案,按照文献[15]中的方法进行编码参数优化,获取这些方案在加权失真下对应的真实比特率,并以除实心点以外的形式展现在图 4 中。分析拟合曲线和其他实际值的对比结果可以看出,通过与序列相关的参数估计,式(17)较好地拟合了划片总数和对应的比特率之间的关系。可以看出,随着划片总数的增加,编码效率开始下降较快,后面趋于平缓。



注:每条曲线上仅有 3 个实心点,表示拟合所用数据

图 4 比特率关于划片数目的拟合曲线

Fig. 4 Fitted curve of bit rate on number of tiles

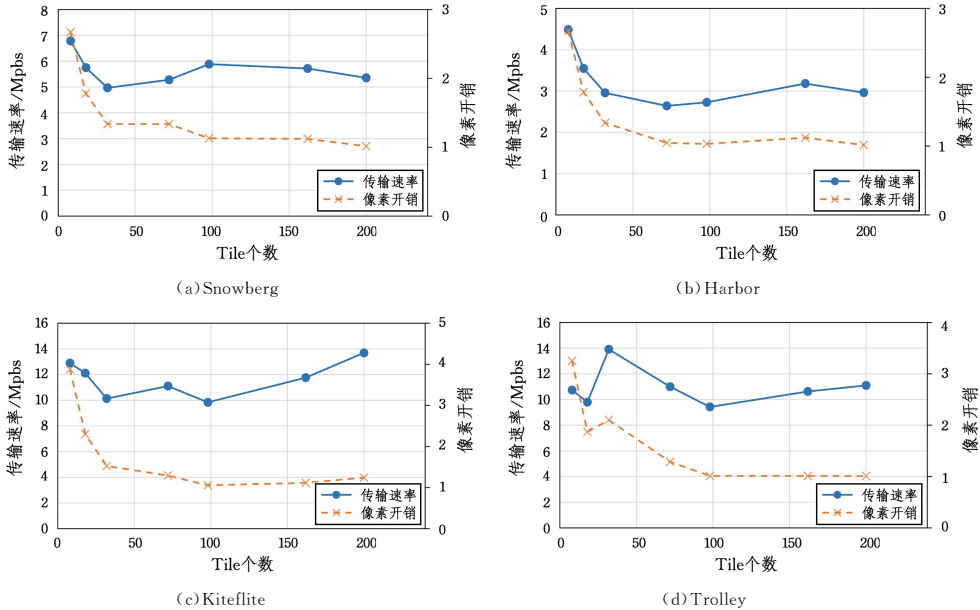


图 5 不同片划分下的像素开销及传输速率

Fig. 5 Pixel overhead and transmission rate with different tile partition schemes

图 5 中的传输速率曲线展示了不同片划分方案下,编码比特率与像素开销乘积即传输速率的变化曲线。将所有可选择的片划分方案代入像素开销和编码效率的计算式,即式(16)和式(17),选择两者乘积最小的方案,即为本文方案得到的片划分方案,如表 3 所列。

为了验证所提方案的有效性,将此方案与片划分穷举搜索算法进行比较。对于表 2 中的每种片划分方案,穷举搜索方案都对测试序列基于 4 个基础量化步长{22, 27, 32, 37}编码 4 次。每个片划分方案下的每次编码均根据文献[15]中提出的量化步长调整方案计算每个片的量化步长,用计算出的量化步长对每个片进行编码、解码,得到视口区域的球型加权失真以及传输比特率。在所有片划分方案中,选取在相同传输比特率下视口失真最小的片划分方案。穷举搜索法得到的结果如表 3 所列。

图 4 中不同序列在相同片划分方案下比特率差异较大,原因是各序列的纹理、运动复杂度不同。除此之外,与普通视频不同的是,全景视频在二维平面上的视口区域变形程度和视点的纬度密切相关,如式(6)和式(7)所示。不同序列的视点位置不一样,导致视口在二维平面上的精确面积有较大差别。由于文献[15]根据片出现在视口区域的不同概率对编码参数进行优化,因此导致了不同序列对应的比特率相差较大。

图 5 给出了不同序列在不同片划分下的像素开销及传输速率。由图 5 可以看出,一般来说,随着划片总数的增加,式(16)中定义的像素开销逐渐减少,传输的总像素与视口的实际像素数更为接近。图 5(c)中划片总数为 32 时比划片总数为 18 时像素开销更高,这是由于对于单个序列来说,视点涉及的片数目不仅与片的尺寸有关,也与视点和片划分的具体位置相关。以 Kiteflite 序列的第一帧为例,划片总数为 18 时共有 8 个 tile 包含了视口区域,而划片总数为 32 时共有 16 个 tile 包含了视口区域,因此划片数目为 32 时对应的像素开销更高。

表 3 穷举搜索算法与所提方案得出的最佳片划分方案

Table 3 Optimal tile partition using the proposed scheme and the exhaustive search algorithm

序列名称	所提方案		穷举算法	
	Tile 个数	Tile 大小	Tile 个数	Tile 大小
Snowberg	32 *	45°×45°*	72	30°×30°
Harbor	72	30°×30°	72	30°×30°
Kiteflite	98	25°×25°	98	25°×25°
Trolley	98	25°×25°	98	25°×25°

注:*表示优化模型最优 tile 选择结果与穷举算法不同

从表 3 中可以看到,本文提出的片划分方案在除了 Snowberg 以外的序列中均得到了与穷举搜索法相同的分片大小。图 6 给出了各种片划分下不同传输速率对应的视口失真。对于 Snowberg 序列,所提方案获得的最佳划分总数为 32,穷举方案获得的最佳划分总数则为 72。由图 6(a)也可以

发现,tile个数为8与18时,在相同传输比特率下视口失真较大,划分总数为72或32时传输性能可以达到最优与次优,Snowberg序列片数目为32与72时表现较为接近。因此,所

提方案在不用对每种片划分编码的情况下达到了与片划分穷举搜索算法相当的效果。图6中不同片划分下传输速率与视口失真的表现也展现了与图5中传输率大体一致的趋势。

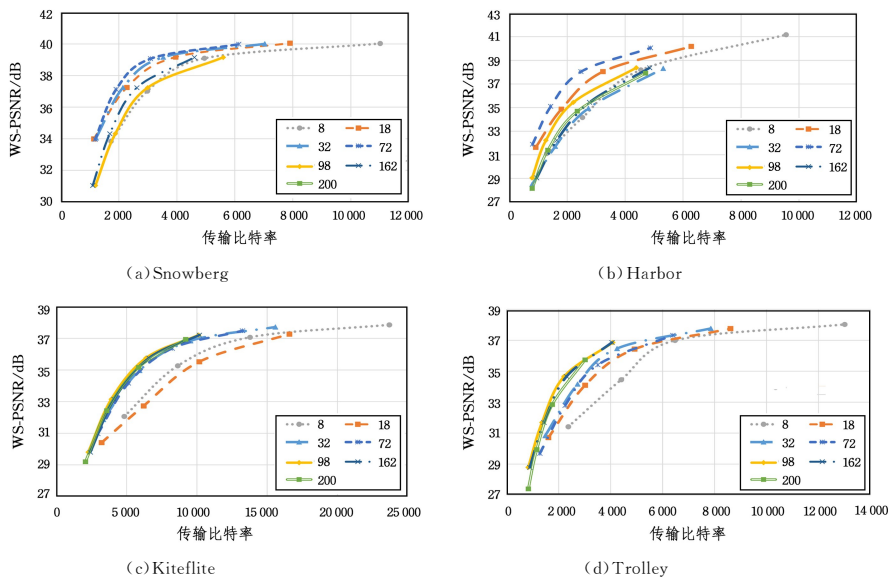


图6 各种片划分方案下不同传输速率对应的视口失真

Fig. 6 Viewport distortion with varying transmission rate for different tile partition schemes

结束语 本文提出了一种ERP投影格式下针对不规则视口的全景视频最优片划分方案。首先,对球面上的视口及二维平面上的不规则视口进行了表示;其次,对不规则视口区域的额外像素传输开销进行了推导;然后,对片编码参数优化的编码方案中切片数目及比特率之间的关系进行了拟合;最后,提出了最佳的片划分方案。将本文方法与片划分穷举搜索算法进行对比,验证了所提的片划分方案的传输性能。

参考文献

- [1] AHMAD A, MANSOOR A, BARAKABITZE A, et al. Supervised-learning-Based QoE Prediction of Video Streaming in Future Networks: A Tutorial with Comparative Study[J]. IEEE Communications Magazine, 2021, 59(11): 88-94.
- [2] ALWIS C D, ANSHUMAN K, PHAM Q V, et al. Survey on 6G Frontiers: Trends, Applications, Requirements, Technologies and Future Research[J]. IEEE Open Journal of the Communications Society, 2021, 2: 836-886.
- [3] NGUYEN D C, MING D, PATHIRANA P N, et al. 6G Internet of Things: A Comprehensive Survey [J]. IEEE Internet of Things Journal, 2022, 9(1): 359-383.
- [4] KAUL A, GUPTA J. Revolutionary 6G: Technologies, Architecture, Coverage, and Performance[C]// Proceedings of International Conference on Computing Communication and Networking Technologies. New York: IEEE Press, 2021: 1-6.
- [5] SKUPIN R, SANCHEZ Y, HELIGE C, et al. Tile Based HEVC Video for Head Mounted Displays[C]// Proceedings of IEEE International Symposium on Multimedia. California: IEEE Press, 2016: 399-400.
- [6] CHIARIOTTI F. A Survey on 360-Degree Video: Coding, Quality of Experience and Streaming[J]. Computer Communications, 2021, 177(1): 133-155.
- [7] SREEDHAR K K, AMINLOU A, HANNUKSELA M M, et al. Viewport-Adaptive Encoding and Streaming of 360-Degree Video for Virtual Reality Applications[C]// IEEE International Symposium on Multimedia. New York: IEEE Press, 2016: 583-586.
- [8] NASRABADI A T, MAHZARI A, BESHAY J D, et al. Adaptive 360-Degree Video Streaming Using Layered Video Coding [C]// IEEE Virtual Reality. California: IEEE Press, 2017: 347-348.
- [9] VISHWANATH B, NANJUNDASWAM T, ROSE K. Rotational Motion Model for Temporal Prediction in 360 video coding[C]// IEEE 19th International Workshop on Multimedia Signal Processing. United Kingdom: IEEE Press, 2017: 1-6.
- [10] FREMEREY S, SINGLA A, MESEBERG K. Avtrack360: An Open Dataset and Software Recording People's Head Rotations Watching 360° Videos on an HMD[C]// ACM Multimedia System Conference. New York: ACM Press, 2018: 403-408.
- [11] XU M, LI C, ZHANG S, et al. State-of-the-Art in 360° Video/Image Processing: Perception, Assessment and Compression[J]. IEEE Journal of Selected Topics in Signal Processing, 2020, 14(1): 5-26.
- [12] 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.
- [13] LI L, YAN N, LI Z, et al. λ -Domain Perceptual Rate Control for 360-Degree Video Compression[J]. IEEE Journal of Selected Topics in Signal Processing, 2020, 14(1): 130-145.
- [14] ZHOU Y, TIAN L, ZHU C, et al. Video Coding Optimization for Virtual Reality 360-Degree Source[J]. IEEE Journal of Selected Topics in Signal Processing, 2020, 14(1): 118-129.

- [15] XU Y, YANG T, TAN Z, et al. FoV-based Coding Optimization for 360-degree Virtual Reality Videos[C]//Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Singapore; IEEE Press, 2022.
- [16] LI Y, XU J, CHEN Z. Spherical Domain Rate-Distortion Optimization for 360-Degree Video Coding[C]//Proceedings of IEEE International Conference on Multimedia and Expo, Hong Kong: IEEE Press, 2017; 709-714.
- [17] LIU Y, GUO H, ZHU C, et al. Spherical Position Dependent Rate-Distortion Optimization for 360-degree Video Coding [C]//Proceedings of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, Lanzhou: IEEE Press, 2019; 992-996.
- [18] CHIANG J, YANG C, DEDHIA B, et al. Saliency-Driven Rate-Distortion Optimization for 360-Degree Image Coding[J]. *Multimedia Tools and Applications*, 2021, 80(6): 8309-8329.
- [19] YANG C, CHIANG J, LIE W. Rate-Distortion Optimization for 360-degree Image Considering Visual Attention[C]//Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, Los Alamitos; IEEE Computer Society, 2020; 1128-1131.
- [20] SUN Y, YU L. Coding Optimization Based on Weighted-to-Spherically-Uniform Quality Metric for 360 Video [C]// Proceedings of IEEE Visual Communications and Image Processing, USA; IEEE Press, 2017; 1-4.
- [21] LUZ G, ASCENSO J, BRITES C, et al. Saliency-Driven Omnidirectional Imaging Adaptive Coding; Modeling and Assessment [C]// Proceedings of IEEE 19th International Workshop on Multimedia Signal Processing, United Kingdom; IEEE Press, 2017; 1-6.
- [22] XIAO M, ZHOU C, LIU Y, et al. Optile: Toward Optimal Tiling in 360-Degree Video Streaming[C]// Proceedings of the 25th ACM International Conference on Multimedia, New York; ACM Press, 2017; 708-716.
- [23] YU M, LAKSHMAN H, GIROD B. Content Adaptive Representations of Omnidirectional Videos for Cinematic Virtual Reality[C]//Proceedings of the 3rd International Workshop on Immersive Media Experiences, New York; ACM Press, 2015; 1-6.
- [24] MAVLANKAR A, GIROD B. Spatial-Random-Access-Enabled Video Coding for Interactive Virtual Pan/Tilt/Zoom Functionality[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2011, 21(5): 577-588.
- [25] GUAN Y, ZHENG C, ZHANG X, et al. Pano: Optimizing 360 Video Streaming with A Better Understanding of Quality Perception [C]//Proceedings of the ACM Special Interest Group on Data Communication, New York; ACM Press, 2019; 394-407.
- [26] SANCHEZ Y, SKUPIN R, HELLGE C, et al. Spatio-Temporal Activity based Tiling for Panorama Streaming[C]//Proceedings of the 27th Workshop on Network and Operating Systems Support for Digital Audio and Video, New York; ACM Press, 2017; 61-66.
- [27] OZCINAR C, CABRERA J, SMOLIC A. Omnidirectional Video Streaming Using Visual Attention-Driven Dynamic Tiling for VR[C]//IEEE Visual Communications and Image Processing, New York; IEEE Press, 2018; 1-4.
- [28] KAN N, ZOU J, LI C, et al. RAPT360: Reinforcement Learning-Based Rate Adaptation For 360-Degree Video Streaming with Adaptive Prediction and Tiling[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, 67(2): 409-423.
- [29] GUO H W, LUO H J, LIU S, et al. Improved R- λ Model Based Rate Control Algorithm [J]. *Computer Science*, 2019, 46(3): 142-147.
- [30] CHEN L, GAO L, REN J, et al. Adaptive Bitrate Streaming for Energy-Efficiency Mobile Augmented Reality [J]. *Computer Science*, 2022, 49(1): 194-203.
- [31] OZCINAR C, CABRERA J, SMOLIC A. Visual Attention-Aware Omnidirectional Video Streaming Using Optimal Tiles for Virtual Reality[J]. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 2019, 9(1): 217-230.
- [32] NGUYEN D V, TRAN H, THANG T. Adaptive Tiling Selection for Viewport Adaptive Streaming of 360-degree Video[J]. *IEICE Transactions on Information and Systems*, 2019, 102(1): 48-51.
- [33] KATTADIGE C, THILAKARATHNA K. VAD360: Viewport Aware Dynamic 360-Degree Video Frame Tiling [J]. arXiv: 2105.11563, 2021.
- [34] JEAN L, CONCOLATO C. Tiled-Based Adaptive Streaming Using MPEG-DASH[J]. *Multimedia Systems*, 2016, 41(3): 1-3.



YANG Tao-yu, born in 1998, master, is a member of China Computer Federation. Her main research interests include immersive videos and video coding.



XU Yuan-yuan, born in 1983, Ph.D, associate professor, postgraduate supervisor, is a member of China Computer Federation. Her main research interests include image/video coding, image/video communication, immersive videos, visual perception and applications.

(责任编辑:喻黎)