摘要

最近虚拟现实技术的突破，特别是在经济型虚拟现实耳机和大规模智能手机方面，正在创造对3D沉浸式虚拟现实内容的快速增长需求。360虚拟现实视频在各个方向记录周围的环境，给用户一个完全沉浸的体验。由于过去几年推出了大量的360摄像头，360视频内容的创建正在迅猛发展，360虚拟现实视频正在成为数字行业的新视频标准。当ERP和CMP可能是存储360个虚拟现实视频最常用的投影和打包布局时，它们具有严重的投影失真、内部不连续接缝或纵横比方面的缺点。我们引入了一种新的格式，使用两阶段映射来打包和存储360个虚拟现实视频。半球是无缝统一映射到正方形上的。将两个各自的正方形缝合成一个长宽比为2:1的矩形。我们的方法能够避免内部的不连续性，产生均匀的像素分布，同时保持高宽比接近16:9的大多数标准高宽比。

CCS概念

•以人为中心的计算→虚拟现实；•计算方法学→虚拟现实；图像表示；

1. 介绍

360虚拟现实视频使用摄像机阵列[CAB16，AGB 16]或旋转摄像机[KDMW17]同时记录各个方向的周围环境。在回放过程中，观察者位于球形屏幕的中心，使用纹理映射技术将360个视频投影在屏幕上。观看360虚拟现实

视频中，观众戴着头戴式显示器（HMD），从球体的中心观看球形屏幕。每个人都可以独立控制观看方向。使用360虚拟现实视频有很多应用，如虚拟现实拍摄、虚拟现实体育直播、虚拟现实巡回赛、社交虚拟现实通信、机器人远程操作等[lxry18、rpac17]

为了支持高效的存储、访问和处理，球形视频需要投影到二维域上。通常，所有现有的投影格式都将球面与简单的平面域（如矩形、正方形或三角形）相关联。投影生成双射映射，将球面的一点映射到平面域上。然后将这些二维基本体（即矩形、正方形或三角形）打包在一起，生成传统的二维视频。在这些投影中，等矩形投影（ERP）和立方映射投影（CMP）在游戏行业中得到了广泛的应用。谷歌的等角立方体模型（eac）[goo17]被建议用来处理投影失真。其他投影技术

包括紧密八面体投影（COHP）[ph03]和紧密二十面体投影（cisp）[fis43]。

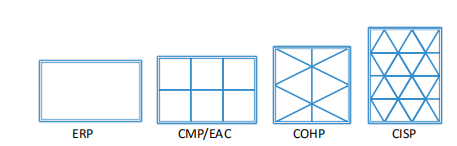


图1：投影和包装布局。

投影和包装布局的三个关键因素影响360虚拟现实视频的质量：像素密度的均匀性、包装边缘的不连续性和矩形纵横比。它们可以用来评估视频质量和存储效率，同时编码360个虚拟现实视频。

像素密度的均匀性：由于投影畸变，一个球体可以在不同的区域以不同的密度采样。例如，erp over对极点处的球体进行采样，从而在二维矩形的顶部和底部产生比其他区域更高的密度。在CMP中，立方体表面的中心更靠近球体，而角则更远离球体。因此，角比中心获得更多像素。EAC通过减少失真来提高像素密度的均匀性。COHP和CISP在映射中使用更多的面，它们更可能生成均匀的像素密度。

包装边缘的不连续性：视频编码标准基于矩形块。因此，在将球面映射到规则多面体上之后，有必要将规则多面体的面打包成二维矩形块。有几种方法可以将这些平面（正方形或三角形）打包成矩形块。图1显示了一些包装布局。但是，在填充这些面时，可能会引入纹理不连续。具有锐边和高对比度的编码帧效率较低，需要特殊处理[MNC99]。尤其是，低保真伪影在高压缩比下更为明显，并且经常限制可以实现的最大压缩。

矩形纵横比：视频压缩对于360个虚拟现实视频至关重要，因为它必须代表一个完整的球形图像，而不仅仅是一个小的视区。例如，为了在给定的观看方向上达到720p的分辨率，整个360虚拟现实视频的分辨率必须大幅提高到6K，这导致带宽需求和流媒体挑战大大增加。立体360虚拟现实视频利用双眼之间的小视差产生三维效果，同时需要双空间来存储左右眼的视频，包装布局会极大地影响视频编码效率。现代视频编码标准基于16:9纵横比[ASP01，SOHW12]。因此，纵横比为16:9或接近16:9的视频在视觉质量和编码效率方面具有优势。

最常见的投影ERP和CMP（或其变体EAC）及其相应的填充布局，显示像素密度变化、内部不连续接缝或纵横比缺点。本文介绍了一种利用连续均匀映射对360个虚拟现实视频进行打包和存储的新格式，当一个半球投影到一个正方形上时，投影保留了邻接和分数投影区域。对于二维单镜VR视频和三维立体VR视频，将两个半球生成的两个正方形组合成一个长宽比为2:1的矩形。我们的方法能够避免内部的不连续性，并产生均匀的像素分布。此外，我们的包装布局的纵横比（即16:8）非常接近大多数标准纵横比16:9。

相关工作

在地图学[sny87]、表面参数化[fh05]、平铺[gs86]和计算机图形学中的纹理映射[wd97，pck04]中对投影和填充布局进行了广泛的研究。在此，我们简要回顾了一些相关工作，并建议参考调查[FH05]了解更多详细信息。

在这些投影中，等矩形投影（ERP）在游戏业和制图学中得到了广泛的应用。此投影将展开一个球体并将其展平为二维矩形。cubemap投影（cmp）[gre86]是另一种常用的方法。它将球体变形为立方体，并将1/6球面展平为立方体的表面。然后它展开立方体的六个面并将它们平放。一个简单的变形是将球体嵌入立方体中，并将球体表面向外投射到立方体的表面上。它提供了比ERP更均匀的像素密度。一些变化，如谷歌的等角立方体地图（eac）[goo17]也被提出。EAC进一步降低了投影失真。CMP和EAC都使用六个正方形进行双目标映射，并且已经被Facebook和谷歌用于360视频行业。其他方法使用更复杂的域进行投影，例如紧凑的八面体投影（COHP）[ph03，ed08]和紧凑的二十面体投影（cisp）[fis43]。COHP和CISP将球面三角形曲面投影到平面三角形域上。COHP使用八个三角形，而CISP使用二十个三角形。由于复杂的双目标映射和大量的内部包装边缘，尚不清楚cisp在实践中是否有用。

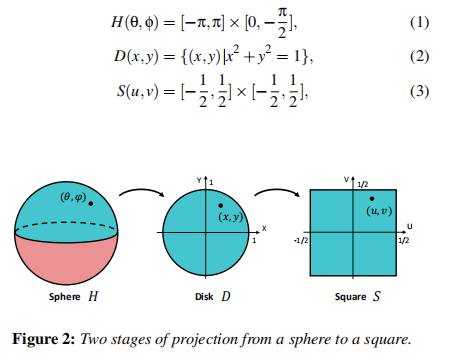
有几种方法可以将这些平面（正方形或三角形）打包成矩形块。一些包装布局如图1所示。纹理不连续

3.概述

我们的目标是在保持投影面积比的同时，在球体和正方形之间建立一个地图。这使我们能够生成具有统一像素分布的高质量360 VR视频。由于缺乏从半球到正方形的明确和直接的映射，我们将映射分为三个阶段。第一阶段执行半球磁盘映射（H→D）。我们用赤道把一个球体分成两个半球。在4.1节中，我们提出了半球和圆盘之间的保留面积比投影。此阶段生成两个磁盘。第二阶段执行disksquare映射（d→s）。提出了圆盘与正方形之间的保持面积比投影（见第4.2节）。这两个映射是双射的，它们的组合提供了一个从球体到正方形（H→D→S）的映射。因为每个阶段都是一个保留的面积比图，所以它们的组成也是如此。这两个映射保证了360个视频的统一像素分布和高效存储。在第三阶段，我们对二维360视频和三维立体360视频应用了不同的策略。对于2d360视频，两个半球的两个正方形被缝合成一个长宽比为16:8的矩形（见第5节）。这种缝合策略保持了边缘之间的原始空间相邻。对于三维立体360视频，第三阶段可分为两个步骤。首先，将两个正方形缝合成一个正方形；然后将左右眼的两个正方形打包，生成一个长宽比为16:8的矩形，接近标准长宽比16:9。

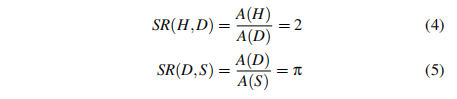
4.保留面积比投影

为了建立一个从球体到正方形的映射，我们使用了两种映射：半球-磁盘映射和磁盘平方映射。给定三个域：一个单位半球h（θ，φ）、一个单位圆盘d（x，y）和一个单位平方s（u，v），我们使用（θ，\_）来表示球面上经度θ和纬度\_的点。（θ，η）对应圆盘上的坐标（x，y）和正方形上的坐标（u，v）。

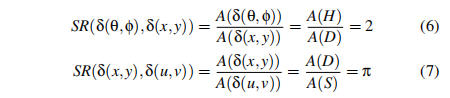


投影的面积比是两个域的分数面积。

两个连续域的面积比定义为其中a（·）是给定域的面积。



如图2所示，对于任一点（θ，φ）h，我们将其邻域表示为δ（θ，φ）h，则其面积为a（δ（θ，φ））。对于对应的地图h→d和d→s，如果分数面积在任意两个对应点的附近保持不变，投影将保留面积比。那就是



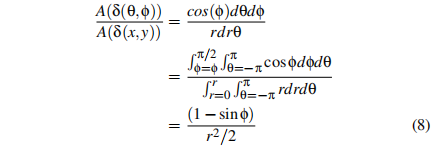
这样可以保证球体上的一组点均匀映射到磁盘上的一组点，而磁盘上的一组点均匀映射到正方形上的一组点。我们将这两个阶段阐述如下。4.1。（θ，φ）∈H和（x，y）∈D附近的半球圆盘投影，它们的面积为

A(δ(θ,φ)) = cos(φ)dθdφ

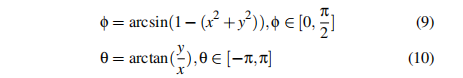
A(δ(x, y)) = rdrdθ

为了简单起见，我们使用（x，y）对应的极坐标（r，θ），其中

把这两个项的分数面积积分，就相当于把一个球形帽映射到一个小圆盘上。对于均匀映射，其分数面积等于整个投影面积比（即2）。



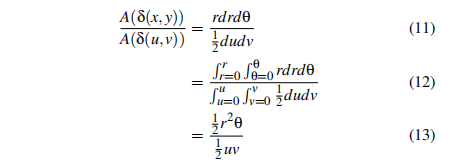
如果投影面积比保持2，我们就得到了从半球到圆盘的下列投影。



这个投影是可逆的。

4.2.圆盘直角投影

给定单位圆盘上的一个点和单位平方上的一个点，我们可以类似地求出它们的分数面积。如等式12所示，通过将公式的两个项在11中进行积分，我们得到圆扇形和直角三角形之间的分数面积（也参见图3）。



这里，如果我们让r=u，得到的分数面积等于整个投影面积比π，我们得到了从一个圆盘到一个正方形的均匀映射。

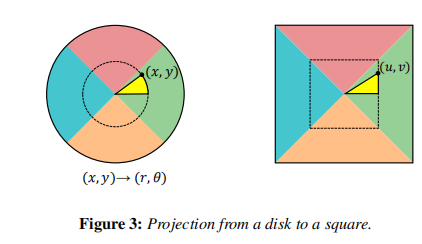


其中，我们只推导出u≥v区域的公式。对于|u| < |v |的区域，其公式可以类似地导出。将极坐标（r，θ）转换回（x，y）后，我们得到了从圆盘到正方形的投影。这个投影是双射的。



从D中的子域到相应子域S的投影面积比可以使用以下雅可比行列式[Rud64]的行列式来计算。因此，我们可以用它的雅可比行列式来验证我们的投影。当|u| >= |v |时，雅可比行列式为





当|u| < |v|时，雅可比的行列式为



这两个雅可比行列式证明了投影面积比是常数。

4.3。饱和度图

我们用饱和度图来显示投影中的像素密度。饱和度图中的编码颜色范围是蓝色、绿色、黄色、橙色和红色。绿色表示球面三角形和平面三角形之间的最佳像素密度比接近1:1。红色表示投影矩形的像素密度不足，而蓝色表示浪费矩形像素，从而破坏纹理。图10显示了一些现有投影和我们的投影的饱和图。我们的投影能够得到完整的云状饱和度图。

在饱和度地图中，颜色在地图中的变化越小，我们可以获得更均匀的像素密度。我们计算了常用投影格式的饱和度图的标准推导（sd），结果如表1所示。我们的方法比其他投影显示出更好的一致性。

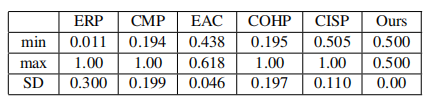
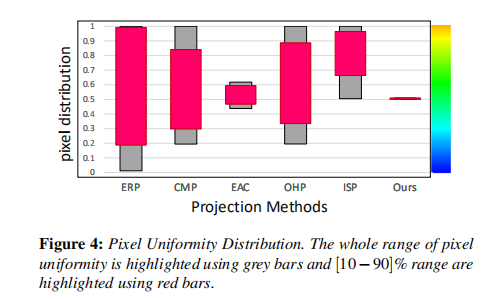


表1:ERP、CMP、EAC、COHP、CISP的投影面积比范围及其标准推导及我们的方法。

在图4中，我们展示了投影面积比的范围，并使用灰色条突出显示它们。因为我们的分数区域是常数，所以我们简单地把它设为0.5。如果我们删除10%的低密度数据和10%的高密度数据，我们将得到用红色范围条突出显示的[10-90]%范围结果。我们的投影在像素均匀性上超过了所有其他方法。

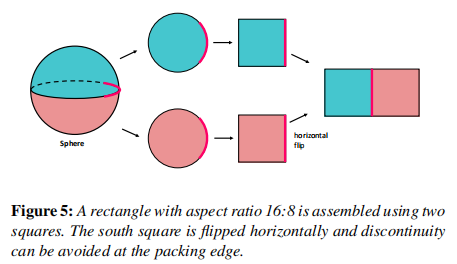
5.包装布局

由于360虚拟现实视频比传统的二维视频需要更多的数据来存储和传输，视频压缩变得非常关键。编码人员经常压缩连续的内容，其保真度比边缘锐利、对比度高的合成图像要好。为了获得良好的编码效果，帧内预测利用相邻像素之间的相关性，从已编码的像素预测未编码像素的值。只有相邻像素突变时才会存储差异[SOHW12]。因此，包装边缘的不连续性会影响编码效率。在这里，我们描述了投影后将布局打包到传统视频上的过程。我们为单镜（2D）和立体（3D）360视频演示了不同的解决方案。此外，我们的包装布局生成的视频的纵横比接近视频行业标准16:9。

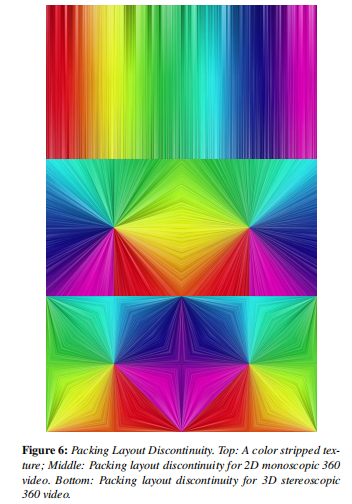


5.1.二维视频打包布局

投影后，分别为北半球和南半球生成两个正方形。它们的四条边都对应于球体的赤道，每条边都对应于1/4的赤道。如图5所示，正方形的两个公共边以红色突出显示。我们把这两个方格并排地放在一起，用一个共同的边把它们连在一起。请注意，我们在缝合前水平翻转南方。然后我们得到一个无缝矩形

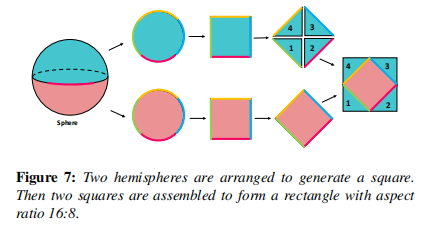


在这里，我们使用一个彩色剥离纹理来说明我们的包装布局的不连续性。由于纹理不连续会导致视频编码中的视觉伪影和效率降低，因此避免相邻封装块之间的不连续变得至关重要。在图6（顶部）中，我们显示了一个彩色的剥离纹理。纹理的上部将映射到北半球，下部将映射到南半球。图6（中间）显示了使用此纹理的球体和填充布局的投影。显然，没有内部不连续和缝合不连续。此外，我们还将该方法与其他投影和包装布局（CMP/EAC、COHP和CISP）进行了比较。更多详情请参考第5.3.1节和图8。



5.2.三维视频打包布局

立体360虚拟现实视频需要为右眼和左眼提供单独的图像。为了满足视频纵横比16:9的要求，我们提出了另一种包装布局。首先，两个正方形都逆时针旋转45度。第二，北广场被分成四个直角三角形。第三，将四个三角形缝合到南广场的角上，生成一个新的大广场，如图7所示。该安排与[ph03，ed08]中建议的安排类似。由于所有原始相邻边都是相邻的，因此内部连续性完全保持在生成的布局中。存储360立体视频时，将双眼的图像并排放置，形成宽高比为2:1（即16:8）的视频。由于左侧图像的右侧与右侧图像的左侧相交，因此包装边缘将有接缝。因此，包装边缘的视频纹理是不连续的。我们建议水平翻转右图像，然后将其与左图像缝合。左图像的右侧与右图像的右侧相交。由于左眼和右眼的图像都有很小的差异，很可能纹理在同一侧保持相似。这有助于避免包装边缘出现明显的接缝和不连续性。

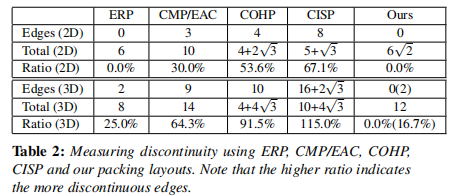


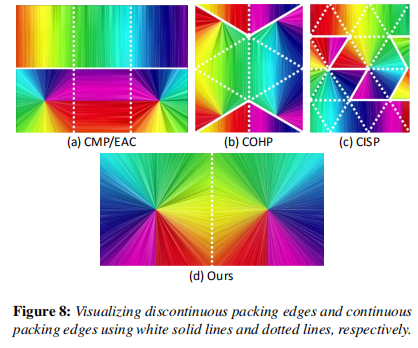
5.3.比较

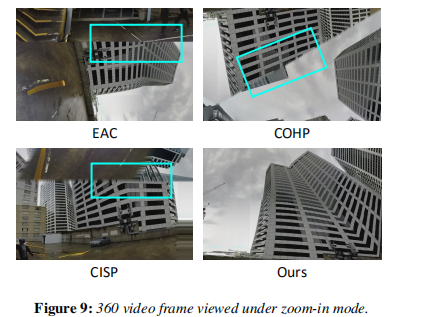
我们进行了实验，以比较我们的方法与其他预测和包装布局。

5.3.1.不连续性比较

如图8所示，使用白色实线突出显示不连续的相邻面，使用白色虚线突出显示连续的相邻面。在图9中，我们放大一个视频帧来演示这些不连续性。我们还使用不连续边缘与总包装边缘的比率定量测量不连续性。结果如表2所示。注意，比率越高，不连续的边越多。我们的方法和ERP都没有不连续的接缝。在这些包装布局中，cisp的连续性最差。对于三维包装布局中的不连续边缘（表2中的最后一列），即使在包装布局有明显接缝的情况下，结果比率（16.7%）仍优于其他方法。





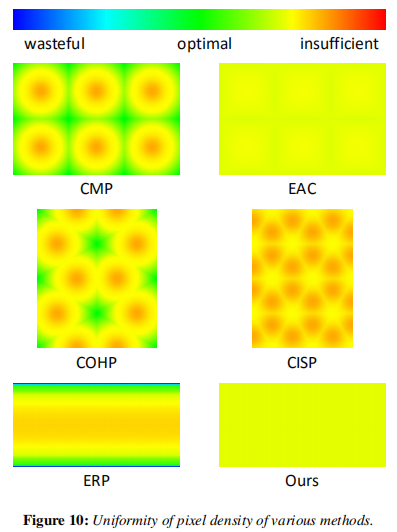


5.3.2.像素均匀性比较

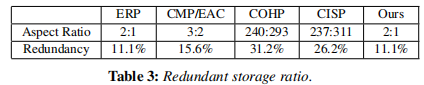
我们还比较了我们的方法与其他投影和包装布局的像素均匀性。如图10所示，我们使用饱和度图来显示像素密度。投影面积比在饱和度图中用颜色编码[GOO17]。在投影和填充之后，我们的方法显示出最佳的均匀性。

5.3.3.宽高比

作为国际标准的数字视频格式，纵横比16:9在录制、处理、压缩、存储和重放过程中提供了最佳的效率。由于360 VR视频必须以传统的二维格式存储，因此保持标准纵横比16:9或接近16:9是有益的。此外，还可以使用一些优化的编解码器技术来进一步提高编码效率。在某些情况下，没有纵横比16:9的视频需要调整为16:9，这会导致存储冗余。我们的包装布局生成了视频



比例接近16:9。表3显示了一些标准包装布局的纵横比及其存储冗余。



6。实施与结果

在本节中，我们将描述方法的实现细节，并重点介绍其性能。

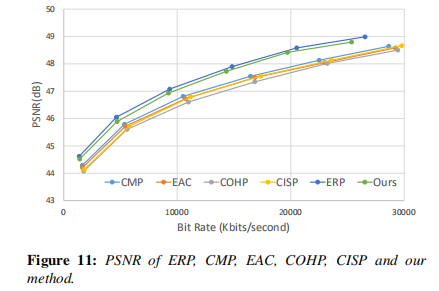
6.1。实施

我们在Windows下用C++语言实现了投影和包装布局。我们使用公共域360视频库360lib来实现我们的算法，并与其他投影进行比较。

6.2。更多结果

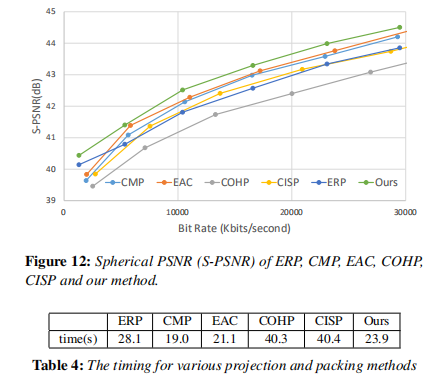
为了进一步证明投影面积比，我们设计了一个球面，其经纬线均匀分布，并且在球面上有几个面积相同的红色圆盘。如图13所示，投影后，CMP/EAC、COHP和CISP都与磁盘在不同区域的面积存在差异。它们的不连续缝可以清楚地观察到。与其他方法相比，我们的结果显示出更好的性能。更多的实验结果可以在补充资料中找到。

文件。从图14到图16，我们展示了三种不同的360视频场景：具有规则和结构化纹理的建筑、具有非结构化纹理的水族馆和具有密集人群的室内会议。6.3。在峰值信噪比方面，我们采用传统的PSNR（峰值信噪比）来估计360虚拟现实视频在编解码过程中的目标质量。一般来说，较高的psnr表示视频质量较高。图11显示了ERP、CMP、EAC、COHP、CISP和我们的方法的PSNR结果。在比较压缩和解压缩视频帧的质量时，ERP具有最高的PSNR。然而，如图8所示，其像素密度的均匀性是最差的。当我们的方法具有完全一致的像素密度时，它也比CMP、EAC、COHP和CISP具有更高的PSNR。我们还使用现代SPSNR（球形PSNR）来评估360个虚拟现实视频的客观质量。s-psnr对查看方向进行平均，并计算所有视区的加权psnr。如图12所示，我们的方法在所有格式中具有最高的SPSNR。



6.4.计算成本

在表4中，我们展示了我们的投影和打包算法对于具有200万像素的给定帧的计算性能（见最后一列）。我们还将该算法与ERP、CMP、EAC、COHP和CISP进行了比较。在一台装有Intel i7 6700处理器和16G内存的PC上测量定时。我们的算法类似于CMP和EAC，但比其他算法更快。



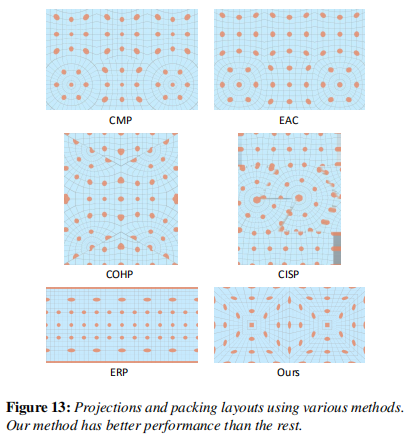
7。结论与未来工作

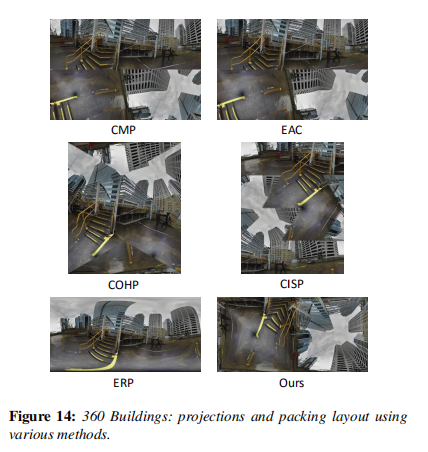
我们提出了一个新的投影和包装布局存储360虚拟现实视频。我们已经证明，利用我们的投影，半球可以无缝、均匀地映射到一个正方形上。我们的方法能够避免内部的不连续性，并产生均匀的像素分布。此外，我们的包装布局的纵横比（即2:1或16:8）非常接近大多数标准纵横比16:9。我们希望我们的方法能大大提高360视频质量，并激励更多的工作沿着这一方向发展。

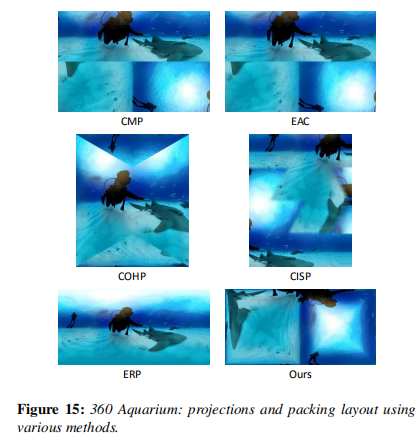
在当前的实现中，我们不执行任何额外的优化。但是，一些优化（如并行计算、边缘填充等）可能会进一步加速整个性能。

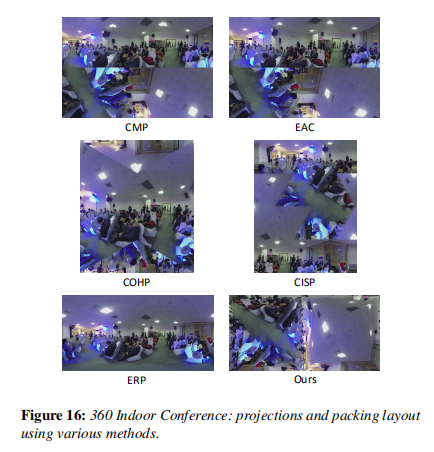
致谢

这项研究工作得到了国家自然科学基金（第61631166002号、第61572196号）的资助。我们要感谢华东师范大学智能机器人运动与视觉实验室和虚拟现实创新实验室的姚昭在算法实现方面的帮助。我们还感谢VR8技术有限公司（中国杭州）的李文松对我们的帮助。









参考文献

[AGB∗16] ANDERSON R., GALLUP D., BARRON J. T., KONTKANEN J., SNAVELY N., HERNÁNDEZ C., AGARWAL S., SEITZ S. M.: Jump: Virtual Reality Video. ACM Transactions on Graphics 35, 6 (Nov. 2016),198:1–198:13. 1

[Asp01] Understanding Aspect Ratios. Technical bulletin, The CinemaSource Press, 2001. 2

[Cab16] CABRAL B. K.: Introducing Facebook Surround 360: An open,high-quality 3D-360 video capture system, 4 2016. 1

[ED08] ENGELHARDT T., DACHSBACHER C.: Octahedron environmentmaps. In VMV (2008), pp. 383–388. 2, 5

[FH05] FLOATER M., HORMANN K.: Advances in Multiresolution for Geometric Modelling. Springer, Berlin, Heidelberg, 2005, ch. Surface Parameterization: a Tutorial and Survey. 2

[Fis43] FISHER I.: A world map on a regular icosahedron by gnomonic projection. Geographical Review 33, 4 (1943), 605–619. 1, 2

[Goo17] GOOGLE: Bringing pixels front and center in vr video,2017. URL: https://blog.google/products/google-vr/bringing-pixels-front-and-center-vr-video/. 1, 2, 6

[Gre86] GREENE N.: Environment mapping and other applications of world projections. IEEE Computer Graphics and Applications 6, 11(1986), 21–29. 1, 2

[GS86] GRÜNBAUM B., SHEPHARD G. C.: Tilings and Patterns. W. H.Freeman & Co., New York, NY, USA, 1986. 2

[KDMW17] KONRAD R., DANSEREAU D. G., MASOOD A., WETZSTEIN G.: SpinVR:Towards live-streaming 3D virtual reality video.ACM Transactions on Graphics 36, 6 (Nov. 2017), 209:1–209:12. 1

[LXRY18] LUO B., XU F., RICHARDT C., YONG J.-H.: Parallax360:Stereoscopic 360 scene representation for head-motion parallax. IEEE Transactions on Visualization and Computer Graphics 24, 4 (2018),1545–1553. 1

[MNC99] MEIER T., NGAN K. N., CREBBIN G.: Reduction of blocking artifacts in image and video coding. IEEE Transactions on Circuits and Systems for Video Technology 9, 3 (Apr. 1999), 490–500. 2

[PCK04] PURNOMO B., COHEN J. D., KUMAR S.: Seamless texture atlases. In Eurograph-ics/SIGGRAPH symposium on Geometry Processing (2004), pp. 65–74.2

[PH03] PRAUN E., HOPPE H.: Spherical parametrization and remeshing.In ACM Transactions on Graphics (2003), vol. 22, ACM, pp. 340–349.1, 2, 5

[RPAC17] RHEE T., PETIKAM L., ALLEN B., CHALMERS A.: Mr360:Mixed reality rendering for 360 panoramic videos. IEEE Transactions on Visualization and Computer Graphics 23, 4 (2017), 1379–1388. 1

[Rud64] RUDIN W.: Principles of mathematical analysis, vol. 3.McGraw-hill New York, 1964. 3

[Sny87] SNYDER J. P.: Map projections–A working manual, vol. 1395.US Government Printing Office, 1987. 2

[Sny97] SNYDER J. P.: Flattening the earth: two thousand years of map projections. University of Chicago Press, 1997. 1, 2

[SOHW12] SULLIVAN G. J., OHM J.-R., HAN W.-J., WIEGAND T.:Overview of the high efficiency video coding (hevc) standard. IEEE Transactions on Circuits and Systems for Video Technology 22, 12 (Dec.2012), 1649–1668. 2, 4

[WD97] WEINHAUS F. M., DEVARAJAN V.: Texture mapping 3D models of real-world scenes. ACM Computer Survey 29, 4 (Dec. 1997), 325–365. 2

[YLG15] YU M., LAKSHMAN H., GIROD B.: A framework to evaluate omnidirectional video coding schemes. In IEEE International Symposium on Mixed and Augmented Reality (2015), pp. 31–36. 7