成 都 理 工 大 学

学士学位论文（设计）外文译文

学生姓名：郑宝仁	学号：201513070304	专业名称：通信工程
译文标题（中英文）：A New Uniform Format for 360 VR Videos/一种新的360虚拟现实视频统一格式		
译文出处： 太平洋图形2018第37卷（2018年），第7期		指导教师审阅**签名**：

外文译文正文：

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，导致它们有严重的投影失真、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。先将半球无缝统一映射到正方形上的，然后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视频内部的不连续性，产生均匀的像素分布，同时保持高宽比接近16:9的大多数标准视频高宽比。

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]。

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。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

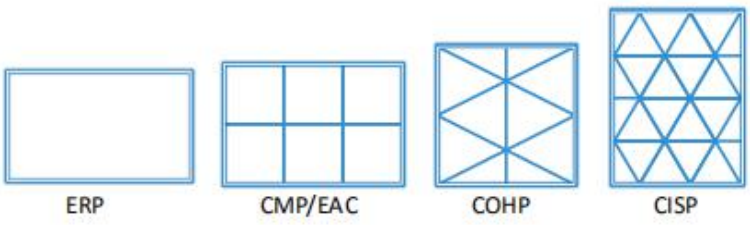

图1：投影和包装布局。

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）。

$$H(\theta,\phi)=[-\pi,\pi]\times[0,-\frac{\pi}{2}], \quad (1)$$

$$D(x,y)=\{(x,y)|x^2+y^2=1\}, \quad (2)$$

$$S(u,v)=[-\frac{1}{2},\frac{1}{2}]\times[-\frac{1}{2},\frac{1}{2}], \quad (3)$$

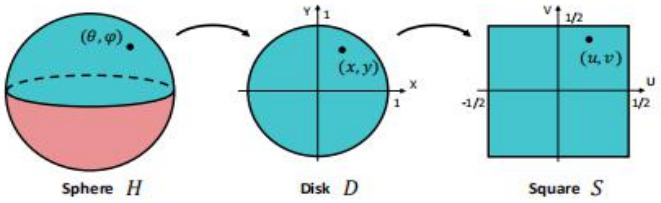

Figure 2: *Two stages of projection from a sphere to a square.*

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。

$$SR(H,D)=\frac{A(H)}{A(D)}=2 \quad (4)$$

$$SR(D,S)=\frac{A(D)}{A(S)}=\pi \quad (5)$$

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

$$SR(\delta(\theta,\phi),\delta(x,y))=\frac{A(\delta(\theta,\phi))}{A(\delta(x,y))}=\frac{A(H)}{A(D)}=2 \quad (6)$$

$$SR(\delta(x,y),\delta(u,v))=\frac{A(\delta(x,y))}{A(\delta(u,v))}=\frac{A(D)}{A(S)}=\pi \quad (7)$$

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

$A(\delta(\theta,\phi))=\cos(\phi)d\theta\,d\phi$

$A(\delta(x,y))=rdrd\theta$

为了简单起见，我们使用（x，y）对应的极坐标（r，θ），其中 $r=\sqrt{x^2+y^2}.$

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）。

$$\frac{A(\delta(\theta,\phi))}{A(\delta(x,y))} = \frac{cos(\phi)d\theta d\phi}{rdr\theta} = \frac{\int_{\phi=\phi}^{\pi/2}\int_{\theta=-\pi}^{\pi}\cos\phi d\phi d\theta}{\int_{r=0}^{r}\int_{\theta=-\pi}^{\pi} rdrd\theta} = \frac{(1-\sin\phi)}{r^2/2} \quad (8)$$

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。

$$\phi = \arcsin(1-(x^2+y^2)), \phi \in [0, \frac{\pi}{2}] \quad (9)$$

$$\theta = \arctan(\frac{y}{x}), \theta \in [-\pi, \pi] \quad (10)$$

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 $(x,y) \in D$ 和单位平方上的一个点 $(x,y) \in D$，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）。

$$\frac{A(\delta(x,y))}{A(\delta(u,v))} = \frac{rdrd\theta}{\frac{1}{2}dudv} \quad (11)$$

$$= \frac{\int_{r=0}^{r}\int_{\theta=0}^{\theta} rdrd\theta}{\int_{u=0}^{u}\int_{v=0}^{v}\frac{1}{2}dudv} \quad (12)$$

$$= \frac{\frac{1}{2}r^2\theta}{\frac{1}{2}uv} \quad (13)$$

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。

$$\begin{cases} r = u, \\ \theta = \pi\frac{u}{v}, \end{cases} \quad (14)$$

其中 $r = \sqrt{x^2+y^2}$，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。

$$(x,y) = \begin{cases} (u\cos(\pi\frac{v}{u}), u\sin(\pi\frac{v}{u})), & |u| \geq |v| \\ (v\sin(\pi\frac{u}{v}), v\cos(\pi\frac{u}{v})), & |u| < |v| \end{cases} \quad (15)$$

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

$$\begin{vmatrix} \cos(\pi\frac{v}{u}) + \pi\frac{v}{u}\sin(\pi\frac{v}{u}) & -\pi\sin(\frac{\pi v}{u}) \\ \sin(\pi\frac{v}{u}) - \pi\frac{v}{u}\cos(\pi\frac{v}{u}) & \pi\cos(\frac{\pi v}{u}) \end{vmatrix} = \pi \quad (16)$$

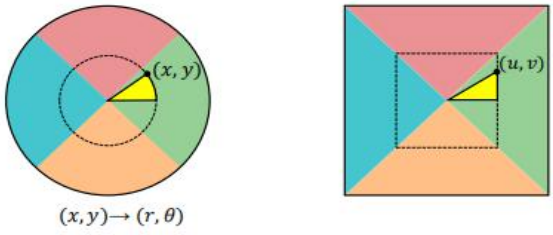

Figure 3: *Projection from a disk to a square.*

当|u| < |v|时，雅可比的行列式为

$$\begin{vmatrix} \pi\cos(\pi\frac{u+v}{v}) & \sin(\pi\frac{u+v}{v}) - \pi\frac{u}{v}\cos(\pi\frac{u+v}{v}) \\ -\pi\sin(\pi\frac{u+v}{v}) & \cos(\pi\frac{u+v}{v}) - \pi\frac{u}{v}\cos(\pi\frac{u+v}{v}) \end{vmatrix} = \pi \quad (17)$$

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。

	ERP	CMP	EAC	COHP	CISP	Ours
min	0.011	0.194	0.438	0.195	0.505	0.500
max	1.00	1.00	0.618	1.00	1.00	0.500
SD	0.300	0.199	0.046	0.197	0.110	0.00

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。

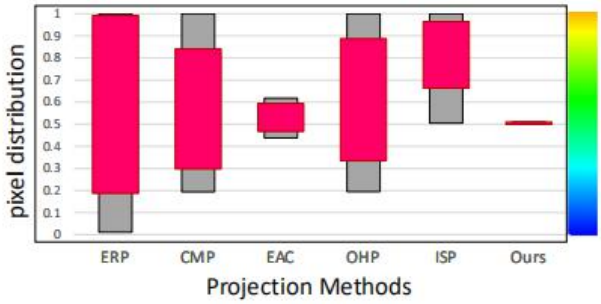

Figure 4: *Pixel Uniformity Distribution. The whole range of pixel uniformity is highlighted using grey bars and* $[10-90]\%$ *range are highlighted using red bars.*

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

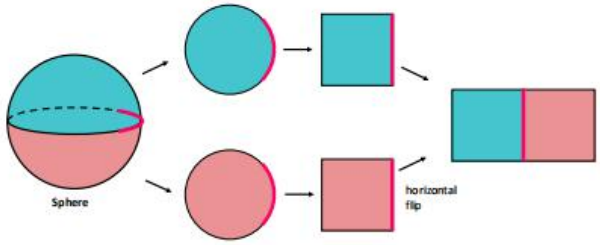

Figure 5: *A rectangle with aspect ratio 16:8 is assembled using two squares. The south square is flipped horizontally and discontinuity can be avoided at the packing edge.*

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。

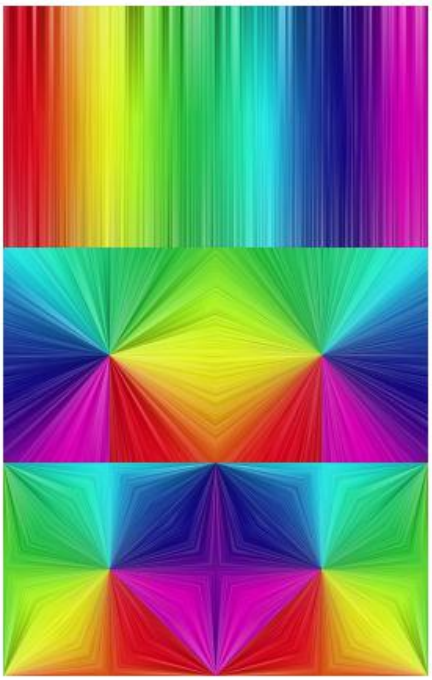

Figure 6: *Packing Layout Discontinuity. Top: A color stripped texture; Middle: Packing layout discontinuity for 2D monoscopic 360 video. Bottom: Packing layout discontinuity for 3D stereoscopic 360 video.*

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。

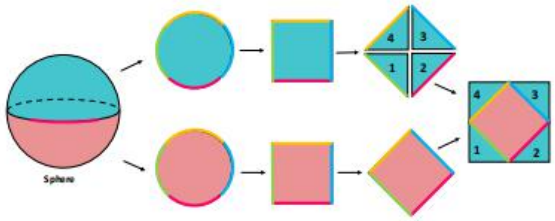

Figure 7: *Two hemispheres are arranged to generate a square. Then two squares are assembled to form a rectangle with aspect ratio 16:8.*

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。

	ERP	CMP/EAC	COHP	CISP	Ours
Edges (2D)	0	3	4	8	0
Total (2D)	6	10	$4+2\sqrt{3}$	$5+\sqrt{3}$	$6\sqrt{2}$
Ratio (2D)	0.0%	30.0%	53.6%	67.1%	0.0%
Edges (3D)	2	9	10	$16+2\sqrt{3}$	0(2)
Total (3D)	8	14	$4+4\sqrt{3}$	$10+4\sqrt{3}$	12
Ratio (3D)	25.0%	64.3%	91.5%	115.0%	0.0%(16.7%)

Table 2: *Measuring discontinuity using ERP, CMP/EAC, COHP, CISP and our packing layouts. Note that the higher ratio indicates the more discontinuous edges.*

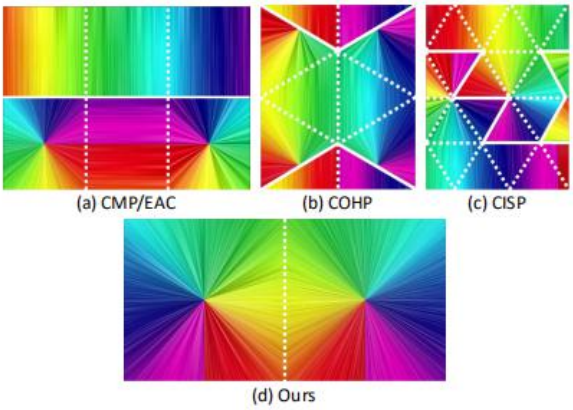

Figure 8: *Visualizing discontinuous packing edges and continuous packing edges using white solid lines and dotted lines, respectively.*

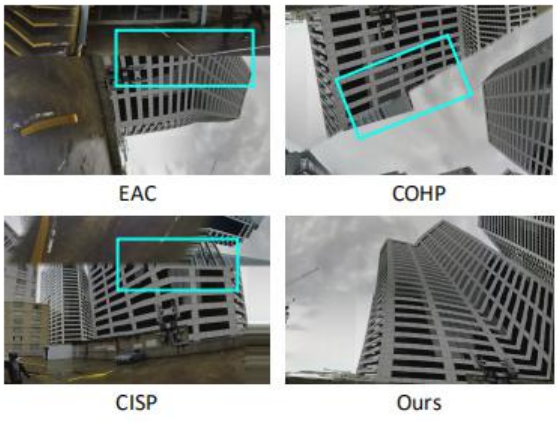

Figure 9: *360 video frame viewed under zoom-in mode.*

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

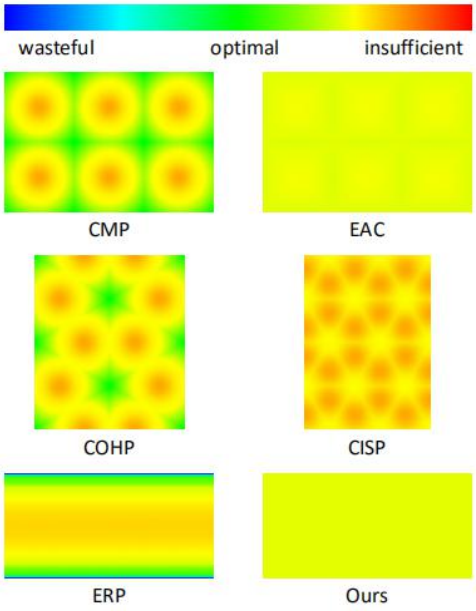

Figure 10: *Uniformity of pixel density of various methods.*

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。

	ERP	CMP/EAC	COHP	CISP	Ours
Aspect Ratio	2:1	3:2	240:293	237:311	2:1
Redundancy	11.1%	15.6%	31.2%	26.2%	11.1%

Table 3: *Redundant storage ratio.*

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。

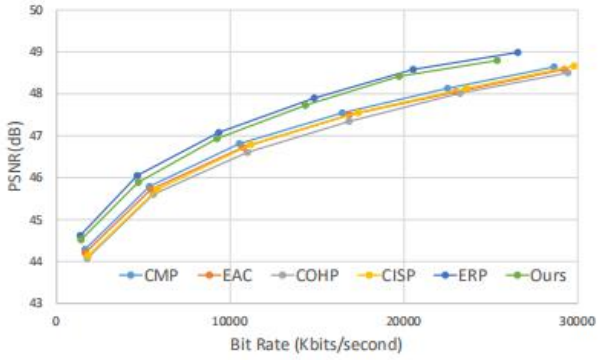

Figure 11: *PSNR of ERP, CMP, EAC, COHP, CISP and our method.*

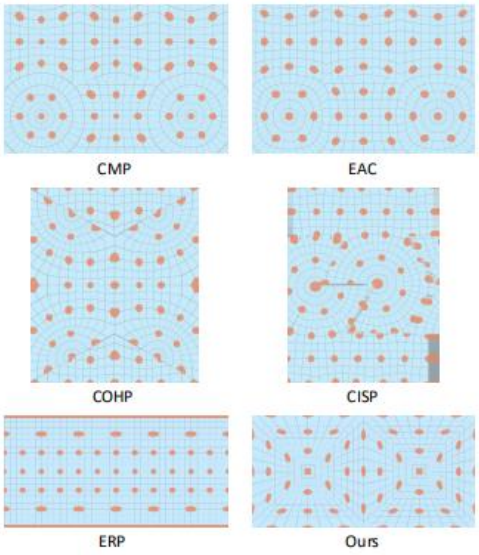

Figure 13: *Projections and packing layouts using various methods. Our method has better performance than the rest.*

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。

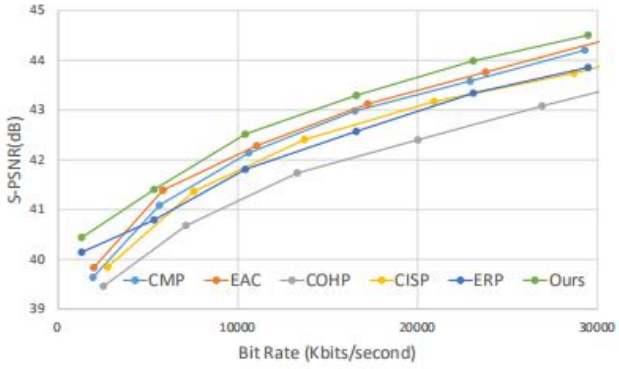

Figure 12: *Spherical PSNR (S-PSNR) of ERP, CMP, EAC, COHP, CISP and our method.*

	ERP	CMP	EAC	COHP	CISP	Ours
time(s)	28.1	19.0	21.1	40.3	40.4	23.9

Table 4: *The timing for various projection and packing methods*

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。

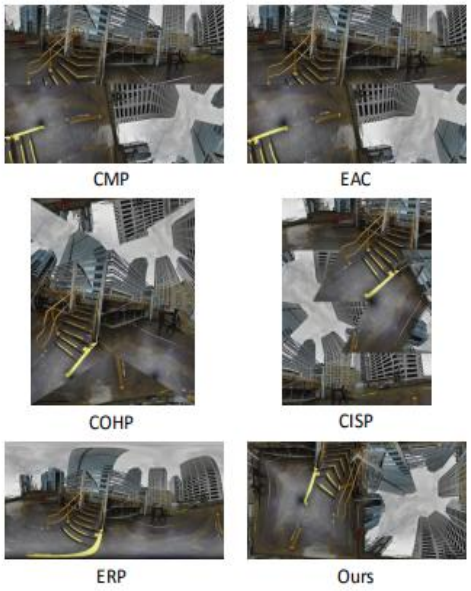

Figure 14: *360 Buildings: projections and packing layout using various methods.*

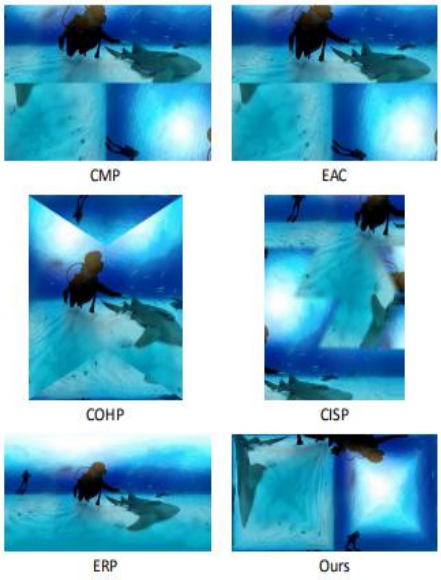

Figure 15: *360 Aquarium: projections and packing layout using various methods.*

Figure 16: *360 Indoor Conference: projections and packing layout using various methods.*

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 Eurographics/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

译文原文：

DOI: 10.1111/cgf.13564

Pacific Graphics 2018 Volume 37 (2018), Number 7

H. Fu, A. Ghosh, and J. Kopf

(Guest Editors)

A New Uniform Format for 360 VR Videos

J. Guo, Q. K. Pei, G. L. Ma, L. Liu and X. Y. Zhang†

School of Computer Science & Software Engineering, East China Normal University, China.

xyzhang@sei.ecnu.edu.cn

Abstract

Recent breakthroughs in VR technologies, especially in economic VR headsets and massive smartphones are creating a fastgrowing demand for 3D immersive VR content. 360 VR videos record a surrounding environment in every direction and give users a fully immersive experience. Thanks to a ton of 360 cameras that launched in the past years, 360 video content creation is exploding and 360 VR videos are becoming a new video standard in the digital industry. When ERP and CMP are perhaps the most prevalent projection and packing layout for storing 360 VR videos, they have severe projection distortion, internal discontinuous seams or disadvantages in aspect ratio. We introduce a new format for packing and storing 360 VR videos using two stage mappings. Hemispheres are seamlessly and uniformly mapped onto squares. Two respective squares are stitched to form a rectangle with the aspect ratio 2 : 1. Our approach is able to avoid internal discontinuity and generate uniform pixel distribution, while keeping the aspect ratio close to the majority standard aspect ratio of 16 : 9.

CCS Concepts

•Human-centered computing → Virtual reality; •Computing methodologies → Virtual reality; Image representations;

1. Introduction

360 VR videos record a surrounding environment in every direction at the same time using a camera array [Cab16, AGB* 16] or a spinning camera [KDMW17]. During playback, the viewer locates at the center of a spherical screen on which 360 videos are projected using texture mapping techniques. To watch a 360 VR

video, a viewer wears a head-mounted display (HMD) and looks at the spherical screen from the center of the sphere. Everyone has independent control of the viewing direction. There are many applications using 360 VR videos such as VR filming, live VR sports broadcasting, VR tour, social VR communication, teleoperation in robotics, etc. [LXRY18,RPAC17]

In order to support efficient storage, access and processing, a spherical video requires to be projected onto a 2D domain. In general, all the existing projection formats associate a spherical surface with simple planar domains like rectangles, squares or triangles. A projection generates a bijective mapping that maps a point of the spherical surface onto a planar domain. These 2D primitives (i.e., rectangles, squares or triangles) are then packed together to generate a traditional 2D video. Among these projections, Equirectangular Projection (ERP) [Sny97] and Cubemap Projection (CMP) [Gre86] are frequently used in gaming industry. Google's Equi-Angular Cubemap (EAC) [Goo17] was suggested to handle projection distortion. Other projection techniques

include Compact Octahedron Projection (COHP) [PH03] and Compact Icosahedron Projection (CISP) [Fis43].

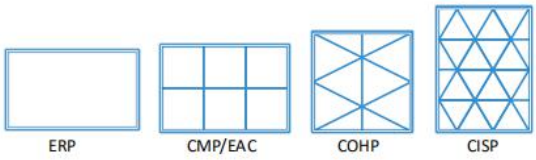

Figure 1: Projections and Packing Layouts.

Three key factors of projection and packing layout affect the quality of 360 VR videos: uniformity of pixel density, discontinuity of packing edges and rectangle aspect ratio. They can be used to evaluate the visual quality and storage efficiency while encoding 360 VR videos.

Uniformity of Pixel Density: Due to projection distortion, a sphere may be sampled with variant density at different regions. For example, ERP over-samples the sphere at poles and results in higher density at the top and bottom of a 2D rectangle than other areas. In CMP, the center of a cube face is closer to the sphere, while the corners are further away. Therefore, the corners get more pixels than the center. EAC improves the uniformity of pixel density by reducing distortion. COHP and CISP use more faces in mapping, and they are more likely generating even pixel density.

Discontinuity of Packing Edges: Video encoding standards are based on rectangular blocks. Therefore, it is necessary to pack the faces of a regular polyhedron into a 2D rectangular block after mapping a spherical surface onto a regular polyhedron. There are a few ways to pack these planar faces (either squares or triangles) into a rectangular block. Figure 1 shows a few packing layout. However, texture discontinuities may be introduced while packing these faces. Encoding frames with sharp edges and high contrast is less efficient and needs special treatment [MNC99]. Especially, the lowfidelity artifacts are more visible at high compression ratios and often limit the maximum compression that can be achieved.

Rectangle Aspect Ratio: Video compression becomes crucial for 360 VR videos, since it must represent a full spherical imagery rather than just a small viewport. For example, in order to achieve 720p

resolution in a given viewing direction, the resolution of the entire 360 VR video must be dramatically increased to 6K. This leads to vastly increasing bandwidth demands and streaming challenge. Stereoscopic 360 VR videos use small disparity between two eyes to generate 3D effects and require double space to store videos for both right and left eyes.The packing layout can greatly influence the video encoding efficiency. Modern video encoding standards are based on the 16:9 aspect ratio [Asp01,SOHW12]. Therefore, a video with the aspect ratio 16:9 or close to 16:9 has the advantage in visual quality and encoding efficiency.

The most prevalent projection ERP and CMP (or its variant EAC) and their corresponding packing layouts, exhibit pixel density variation, internal discontinuous seams or disadvantages in aspect ratio. In this paper, we introduce a new format for packing and storing 360 VR videos using continuous and uniform mapping.When a hemisphere is projected onto a square, the projection preserves adjacency and fractional projection area. For 2D monoscopic VR videos and 3D stereoscopic VR videos, two squares generated from two hemispheres, are assembled to form a rectangle with an aspect ratio 2:1. Our approach is able to avoid internal discontinuity and generate uniform pixel distribution. Moreover, the aspect ratio of our packing layout (i.e., 16:8) is very close to the majority standard aspect ratio 16:9.

Related Work

Projection and packing layout are extensively studied in cartography [Sny87], surface parameterization [FH05], tilling [GS86] and texture mapping in computer graphics [WD97, PCK04]. Here, we briefly review some related work and we suggest referring to the survey [FH05] for more details.

Among these projections, Equirectangular Projection (ERP) [Sny97] has been widely used in gaming industry and cartography. This projection unfolds a sphere and flattens it into a 2D rectangle. Cubemap Projection (CMP) [Gre86] is another frequently used method. It deforms a sphere into a cube and flattens a 1/6 spherical surface into a cube’s face. Then it unfolds the cube’s six faces and packs them flat. A straightforward deformation is to embed the sphere in a cube and project theb spherical surface outwards onto the surface of the cube. It offers better uniformity of pixel density than ERP. A few variations like Google’s Equi-Angular Cubemap (EAC) [Goo17] were also suggested. EAC further reduces projection distortion. Both CMP and EAC use six squares for bijective mapping and have been used in 360 video industry by Facebook and Google. Other methods use more complex domains for projections, such as Compact Octahedron Projection (COHP) [PH03, ED08] and Compact Icosahedron Projection (CISP) [Fis43]. COHP and CISP project spherical triangular surfaces onto planar triangular domains. COHP uses eight triangles and CISP uses twenty triangles. Due ton the sophisticated bijective mapping and a large number of internal packing edges, it is not clear if CISP is useful in practice.

There are a few ways to pack these planar faces (either squares or triangles) into a rectangular block. Some packing layouts are shown in Figure 1. Texture discontinuities may be introduced in these packed rectangles. Especially, the triangle faces in COHP and CISP need to be packed carefully in order to minimize discontinuity between neighboring faces and to improve encoding efficiency. Except ERP, all other packing layouts have discontinuous packing edges, as shown Figure 1.

3. Overview

Our goal is to build a map between a sphere and a square while preserving projection-area-ratios. This allows us to generate high quality 360 VR videos with uniform pixel distribution. Since there lacks explicit and direct mapping from a hemisphere to square, we divide our mapping into three stages. The first stage carries out hemisphere-disk mapping (H → D). We use the equator to divide a sphere into two hemispheres. In Section 4.1, we present a preserving-area-ratio projection between a hemisphere and a disk. This stage generates two disks. The second stage performs disksquare mapping (D → S). A preserving-area-ratio projection between a disk and a square is proposed (see Section 4.2). These two mappings are bijective, and their composition provides a map from a sphere to a square (H → D → S). Since each stage is a preserving-area-ratio map, so is their composition. These two mappings guarantee a uniform pixel distribution and efficient storage for 360 videos. In the third stage, we apply different strategies for 2D 360 videos and 3D stereoscopic 360 videos. For 2D 360 videos, two squares from two hemispheres are stitched to form a rectangle with the aspect ratio 16:8 (see Section 5). This stitch strategy remains the original spacial adjacency between edges. For 3D stereoscopic 360 videos, the third stage can be divided into two steps. First, two squares are stitched to form a single square; then the two squares from right and left eyes are packed to generate a rectangle with the aspect ratio 16:8, which is close to the standard aspect ratio 16:9.

4. Preserving-Area-Ratio Projection

In order to build a map from a sphere to a square, we use a twostage mapping: hemisphere-disk mapping and disk-square mapping. Given three domains: a unit hemisphere H(θ , ϕ), a unit disc D(x, y) and a unit square S(u, v), we use (θ ,ϕ) to denote a point with longitude θ and latitude ϕ on a spherical surface. (θ ,ϕ) correspond to the coordinates (x, y) on the disk and the coordinates (u, v)on the square.

$$H(\theta,\phi) = [-\pi,\pi] \times [0,-\frac{\pi}{2}], \quad (1)$$

$$D(x,y) = \{(x,y)|x^2+y^2 = 1\}, \quad (2)$$

$$S(u,v) = [-\frac{1}{2},\frac{1}{2}] \times [-\frac{1}{2},\frac{1}{2}], \quad (3)$$

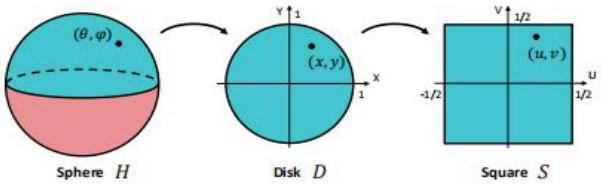

Figure 2: *Two stages of projection from a sphere to a square.*

The area ratio of a projection is a fractional area of two domains.

The area ratios of two successive domains are defined by where A(•) is the area of a given domain.

$$SR(H,D) = \frac{A(H)}{A(D)} = 2 \quad (4)$$

$$SR(D,S) = \frac{A(D)}{A(S)} = \pi \quad (5)$$

As shown in Figure 2, for any point (θ , ϕ) ∈ H, we denote its neighbourhood as δ (θ , ϕ) ⊂ H. Then its area is A(δ (θ , ϕ)). For the corresponding maps H → D and D → S, a projection preserves the area ratio if the fractional area remains constant for the neighbourhood of any two corresponding points. That is

$$SR(\delta(\theta,\phi),\delta(x,y)) = \frac{A(\delta(\theta,\phi))}{A(\delta(x,y))} = \frac{A(H)}{A(D)} = 2 \qquad (6)$$

$$SR(\delta(x,y),\delta(u,v)) = \frac{A(\delta(x,y))}{A(\delta(u,v))} = \frac{A(D)}{A(S)} = \pi \qquad (7)$$

This guarantees that a set of points on the sphere uniformly map to a set on the disk, and a set of points on the disk uniformly map to a set on the square. We elaborate these two stages as follows.4.1. Hemisphere-Disk Projection For the neighborhood of (θ , ϕ) ∈ H and (x, y) ∈ D, their areas are

A(δ (θ , ϕ)) = cos(ϕ)d θ d ϕ

A(δ (x, y)) = rdrd θ

Here, for simplicity, we use the corresponding polar coordinates (r, θ) of (x, y), where $r = \sqrt{x^2+y^2}$.

Integrating both terms of their fractional area is equivalent to mapping a spherical cap to a small disk. For a uniform mapping, their fractional area is equivalent to the entire projection-area-ratio (i.e. 2).

$$\begin{aligned}\frac{A(\delta(\theta,\phi))}{A(\delta(x,y))} &= \frac{cos(\phi)d\theta d\phi}{rdr\theta} \\ &= \frac{\int_{\phi=\phi}^{\pi/2}\int_{\theta=-\pi}^{\pi}\cos\phi d\phi d\theta}{\int_{r=0}^{r}\int_{\theta=-\pi}^{\pi} rdrd\theta} \\ &= \frac{(1-\sin\phi)}{r^2/2} \qquad (8)\end{aligned}$$

If the projection-area-ratio remains 2, we have $\frac{(1-\sin\phi)}{r^2/2} = 2.$ Then we obtain the following projection from a hemisphere to a disk.

$$\phi = \arcsin(1-(x^2+y^2)), \phi \in [0, \frac{\pi}{2}] \qquad (9)$$

$$\theta = \arctan(\frac{y}{x}), \theta \in [-\pi, \pi] \qquad (10)$$

This projection is invertible.

4.2. Disk-Square Projection

Given a point on a unit disk $(x,y) \in D$ and a point on a unit square $(x,y) \in D$, we can analogously derive their fractional area. As shown in Eq.12, by integrating both terms of the formula in 11, we obtain the fractional area between a circular sector and a right triangle (also see Figure 3).

$$\frac{A(\delta(x,y))}{A(\delta(u,v))} = \frac{rdrd\theta}{\frac{1}{2}dudv} \tag{11}$$

$$= \frac{\int_{r=0}^{r}\int_{\theta=0}^{\theta} rdrd\theta}{\int_{u=0}^{u}\int_{v=0}^{v}\frac{1}{2}dudv} \tag{12}$$

$$= \frac{\frac{1}{2}r^2\theta}{\frac{1}{2}uv} \tag{13}$$

Here, if we let r = u and let the resulting fractional area be the same as the entire projection-area-ratio π, we have the following uniform mapping from a disk to a square

$$\begin{cases} r = u, \\ \theta = \pi\frac{u}{v}, \end{cases} \tag{14}$$

where $r = \sqrt{x^2+y^2}$. Note that we only derive the formula for the region where $|u| \geq |v|$. For the region where $|u| < |v|$, its formula can be analogously derived.

After converting polar coordinates (r, θ) back to (x, y), we have the following projection from a disk to a square. This projection is bijective.

$$(x,y) = \begin{cases} (u\cos(\pi\frac{v}{u}), u\sin(\pi\frac{v}{u})), & |u| \geq |v| \\ (v\sin(\pi\frac{u}{v}), v\cos(\pi\frac{u}{v})), & |u| < |v| \end{cases} \tag{15}$$

The projection-area-ratio from a sub-domain in D to a corresponding subdomain S can be calculated using the determinant of the following Jacobian [Rud64]. Therefore, we can verify our projection using the determinant of its Jacobian. When $|u| \geq |v|$, the determinant of Jacobian is

$$\begin{vmatrix} \cos(\pi\frac{v}{u}) + \pi\frac{v}{u}\sin(\pi\frac{v}{u}) & -\pi\sin(\frac{\pi v}{u}) \\ \sin(\pi\frac{v}{u}) - \pi\frac{v}{u}\cos(\pi\frac{v}{u}) & \pi\cos(\frac{\pi v}{u}) \end{vmatrix} = \pi \tag{16}$$

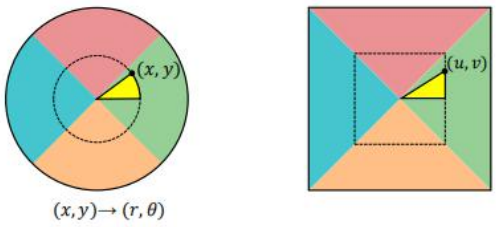

Figure 3: *Projection from a disk to a square.*

When $|u| < |v|$, the determinant of Jacobian is

$$\begin{vmatrix} \pi\cos(\pi\frac{u+v}{v}) & \sin(\pi\frac{u+v}{v}) - \pi\frac{u}{v}\cos(\pi\frac{u+v}{v}) \\ -\pi\sin(\pi\frac{u+v}{v}) & \cos(\pi\frac{u+v}{v}) - \pi\frac{u}{v}\cos(\pi\frac{u+v}{v}) \end{vmatrix} = \pi \tag{17}$$

These two determinants of Jacobian verify that the projection-arearatio is constant.

4.3. Saturation Map

We use a saturation map to display pixel density in projection. Thecoding color in a saturation map ranges through blue, green, yellow, orange, and red. Green indicates an optimal pixel density ratio close to

1:1 between a spherical triangle and a planar triangle.Red indicate insufficient pixel density being projected the rectangle, whereas blue indicates profligate of the rectangle pixels whilestoring texture. Figure 10 shows saturation map of some existingprojections and ours. Our projection is able to achieve completelyuniform saturation map.

In a saturation map, the less the color varies through the map, themore uniform pixel density we can obtain. We compute the standard derivation (SD) for the saturation maps of the common projection formats and the results are shown in Table 1. Our methodexhibits better uniformity than other projection

	ERP	CMP	EAC	COHP	CISP	Ours
min	0.011	0.194	0.438	0.195	0.505	0.500
max	1.00	1.00	0.618	1.00	1.00	0.500
SD	0.300	0.199	0.046	0.197	0.110	0.00

Table 1: The range of projection-area-ratio and its standard derivation for ERP, CMP, EAC, COHP, CISP and our method.

In Figure 4, we show the ranges of projection-area-ratio and highlight them using grey bars. Since our fractional areas are constant, we simply let it be 0.5. If we remove 10% low density data and 10% high density data, we obtain the [10−90]% range results highlighted by red range bars. Our projection surpasses all other methods in pixel uniformity.

5. Packing Layout

Since 360 VR videos require much more data than traditional 2D video to store and transfer, video compression becomes very crucial. A coder often compresses continuous content with better fidelity than synthetic images with sharp edges and high contrast. In order to achieve good encoding results, intra-frame prediction utilizes the correlation between adjacent pixels and predicts the value of uncoded pixels from already coded pixels. The difference is stored only when adjacent pixels are abrupt [SOHW12]. Therefore, discontinuity of packing edges affect encoding efficiency. Here, we describe the process of packing layouts onto traditional videos after projection. We demonstrate different solutions for both monoscopic (2D) and stereoscopic (3D) 360 videos. Moreover, our packing layout generates videos with an aspect ratio close to the video industry standard 16:9

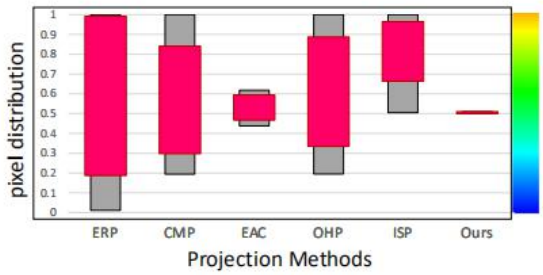

Figure 4: *Pixel Uniformity Distribution. The whole range of pixel uniformity is highlighted using grey bars and* $[10-90]\%$ *range are highlighted using red bars.*

5.1. Packing Layout for 2D Videos

After projection, two squares are generated for the north and south hemisphere, respectively. Both

their four edges correspond to the sphere’ s equator and each edge corresponds to 1/4 equator. As shown in Figure 5, two common edges of the squares are highlighted in red. We stitch the two squares by placing them side by side and adjoining them by a common edge. Note that we horizontally flip the south square before stitching. Then we obtain a seamless rectangle

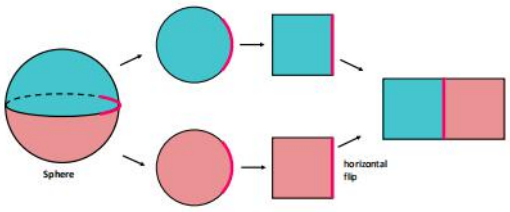

Figure 5: *A rectangle with aspect ratio 16:8 is assembled using two squares. The south square is flipped horizontally and discontinuity can be avoided at the packing edge.*

Here, we use a color stripped texture to illustrate discontinuity in our packing layout. Since texture discontinuities can lead to visual artifacts and less efficiency in video encoding, it becomes crucial to avoid the discontinuity between adjacent packing blocks. In Figure 6-(top), we show a color stripped texture. The upper part of the texture will be mapped to the north hemisphere and lower part will be mapped to the south hemisphere. Figure 6-(middle) shows the projection from a sphere and packing layout using this texture. Obviously, there are no internal discontinuities and no stitching discontinuities. In addition, we compare our method with other projections and packing layouts (CMP/EAC, COHP and CISP). Please refer to Section 5.3.1 and Figure 8 for more details.

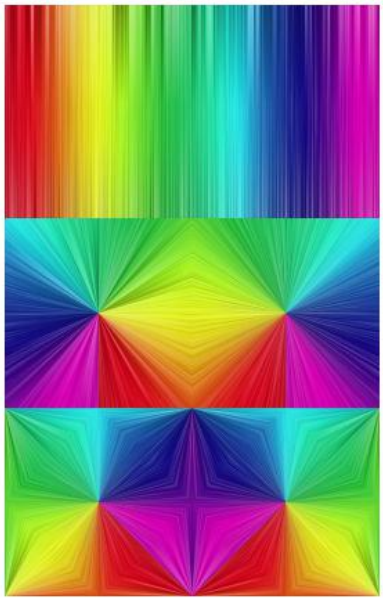

Figure 6: *Packing Layout Discontinuity. Top: A color stripped texture; Middle: Packing layout discontinuity for 2D monoscopic 360 video. Bottom: Packing layout discontinuity for 3D stereoscopic 360 video.*

5.2. Packing Layout for 3D Videos

Stereoscopic 360 VR videos require individual images for both right and left eyes. To meet the requirement of the video aspect ratio 16:9, we propose another packing layout. First, both two squares rotate 45◦ counterclockwise. Second, the north square is split into four right-angled triangles. Third, the four

triangles are stitched to the corners of the south square and generate a new big square as shown in Figures 7. The arrangement is similar to the one suggested in [PH03, ED08]. The internal continuity perfectly remains in the resulting layout since all original neighboring edges are adjacent. When storing a stereoscopic 360 video, the images for both eyes are placed side by side to form a video with the aspect ratio 2:1 (i.e.,16:8). There will be a seam at the packing edge since the right side of the left image meets the left side of the right image. Therefore, the video textures at the packing edges are discontinuous. We propose to horizontally flip the right image and then stitch it with the left image. The right side of the left image meets the right side of the right image. Since the images of both left and right eyes have small disparity, it is very likely that the textures remain similar on the same side. This helps avoid obvious seams and discontinuities at packing edges

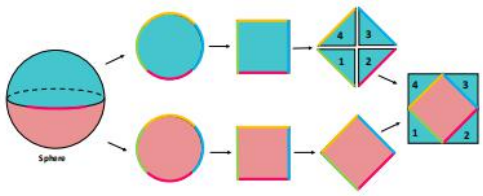

Figure 7: *Two hemispheres are arranged to generate a square. Then two squares are assembled to form a rectangle with aspect ratio 16:8.*

5.3. Comparisons

Experiments were carried out to compare our method with other projections and packing layouts.

5.3.1. Discontinuity Comparison

As shown in Figure 8, the discontinuous adjacent faces are highlighted using white solid lines and the continuous adjacent faces are highlighted using white dotted lines. In Figure 9, we zoom in a video frame to demonstrate these discontinuities. We also quantitatively measure the discontinuity using the ratio of the discontinuous edges to the total packing edges. The results are shown in Table 2. Note that the higher the ratio, the more there are discontinuous edges. Both our method and ERP have no discontinuity seams. Among these packing layouts, CISP has the worst continuity. For the discontinuous edges in our 3D packing layout (the last columns in Table 2), even in the worse case where the packing layout has obvious seam, the resulting ratio (16.7%) is still better than other methods

	ERP	CMP/EAC	COHP	CISP	Ours
Edges (2D)	0	3	4	8	0
Total (2D)	6	10	$4+2\sqrt{3}$	$5+\sqrt{3}$	$6\sqrt{2}$
Ratio (2D)	0.0%	30.0%	53.6%	67.1%	0.0%
Edges (3D)	2	9	10	$16+2\sqrt{3}$	0(2)
Total (3D)	8	14	$4+4\sqrt{3}$	$10+4\sqrt{3}$	12
Ratio (3D)	25.0%	64.3%	91.5%	115.0%	0.0%(16.7%)

Table 2: *Measuring discontinuity using ERP, CMP/EAC, COHP, CISP and our packing layouts. Note that the higher ratio indicates the more discontinuous edges.*

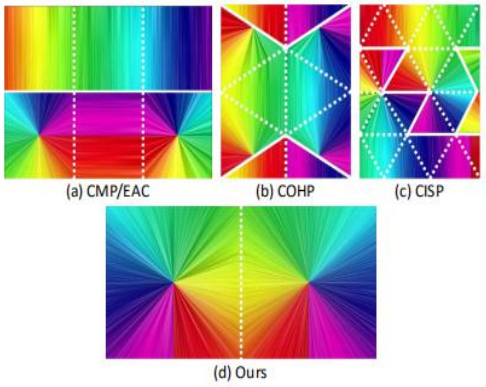

Figure 8: *Visualizing discontinuous packing edges and continuous packing edges using white solid lines and dotted lines, respectively.*

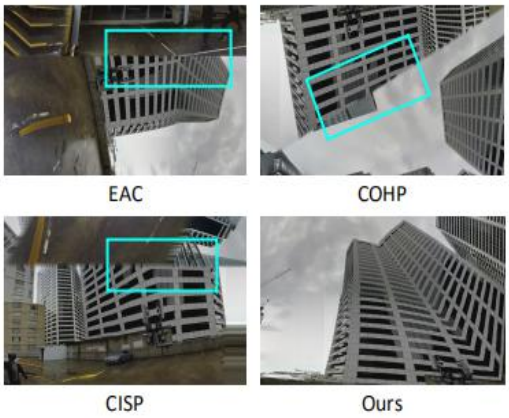

Figure 9: *360 video frame viewed under zoom-in mode.*

5.3.2. Pixel Uniformity Comparison

We also compare our method with other projections and packing layouts in terms of pixel uniformity. As shown in Figure 10, we use the saturation maps to display pixel density. The projection-arearatio is color coded in the saturation map [Goo17]. After projection and packing, our method shows the best uniformity.

5.3.3. Aspect Ratio

As an international standard format of digital videos, the aspect ratio 16:9 offers optimal efficiency during recording, processing, compressing, storage and replaying. Since 360 VR videos have to be stored as traditional 2D formats, it is beneficial to maintain the standard aspect ratio 16:9 or the one close to 16:9. Moreover, some optimized codec techniques can be used to further improve encoding efficiency. In some cases, a video that does not have the aspect ratio 16:9 needs to be resized to 16:9, which results in the storage redundancy. Our packing layout generates videos with the aspect

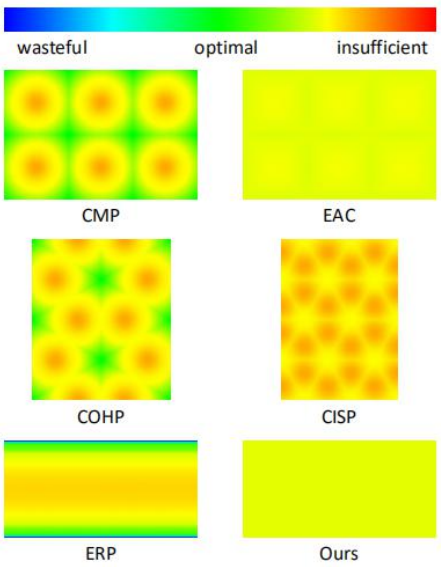

Figure 10: *Uniformity of pixel density of various methods.*

ratio close to 16:9. Table 3 shows the aspect ratio of a few standard packing layouts and their storage redundancy.

	ERP	CMP/EAC	COHP	CISP	Ours
Aspect Ratio	2:1	3:2	240:293	237:311	2:1
Redundancy	11.1%	15.6%	31.2%	26.2%	11.1%

Table 3: *Redundant storage ratio.*

6. Implementation & Results

In this section, we describe the implementation details of our approach and highlight its performance.

6.1. Implementation

We have implemented our projection and packing layout using C++ language under Windows. We use the public domain 360 video library, 360lib , to implement our algorithm and to compare with other projections.

6.2. More Results

In order to further demonstrate projection area ratios, we design a spherical surface where its latitude and longitude lines are evenly distributed, and there are a few red disks with the same area spreading on the surface. As shown in Figure 13, after projection, CMP/EAC, COHP, and CISP all have variance with disk' s areas at different regions. Their discontinuous seams can be clearly observed. Our results exhibit better performance against other methods. More experimental results can be found in the supplemental

document. From Figure 14 to Figure 16, we show three different 360 video scenarios: buildings with regular and structured textures, aquarium with unstructured texture and indoor conference with dense crowd. 6.3. Peak Signal to Noise Ratio Here, we use traditional PSNR(Peak Signal-to-Noise Ratio) to estimate the objective quality of 360 VR videos during the codec process. In general, a higher PSNR indicates that higher quality of a video. Figure 11 shows the results of PSNR for ERP, CMP, EAC, COHP, CISP and our method. ERP has the highest PSNR when comparing the quality of compressed and decompressed video frames. However, as shown in Figure 8, its uniformity of pixel density is the worst. When our method has completely uniform pixel density, it also has higher PSNR than CMP, EAC, COHP and CISP. We also use modern SPSNR(Spherical PSNR) [YLG15] to estimate the objective quality of 360 VR videos. The S-PSNR averages the viewing directions and computes a weighted PSNR over all the viewports. As shown in Figure 12, our method has the highest SPSNR among all the formats.

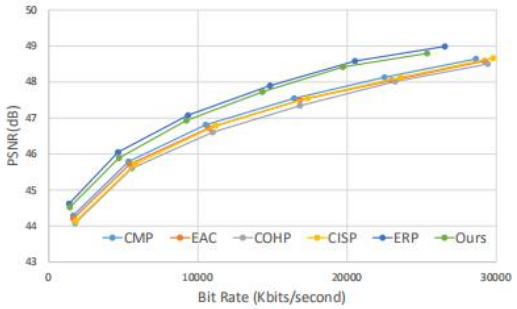

Figure 11: *PSNR of ERP, CMP, EAC, COHP, CISP and our method.*

6.4. Computation Costs

In Table 4, we show the computational performance of our projection and packing algorithm for a

given frame with 2 Megapixels (see the last column). We also compare our algorithm with ERP, CMP, EAC, COHP and CISP. The timing was measured on a PC with Intel i7 6700 processor and 16G of memory. Our algorithms are similar to CMP and EAC, but faster than the rest.

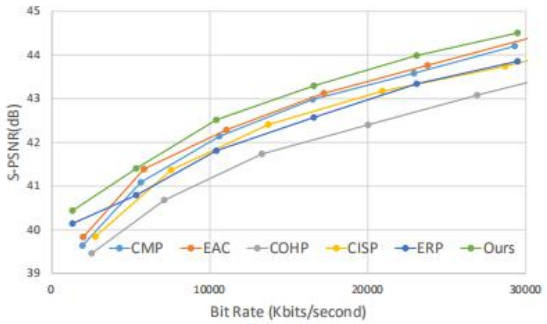

Figure 12: *Spherical PSNR (S-PSNR) of ERP, CMP, EAC, COHP, CISP and our method.*

	ERP	CMP	EAC	COHP	CISP	Ours
time(s)	28.1	19.0	21.1	40.3	40.4	23.9

Table 4: *The timing for various projection and packing methods*

7. Conclusions and Future Work

We present a new projection and packing layout for storing 360 VR videos. We have demonstrated that a hemisphere can seamlessly and uniformly mapped to a square using our projection. Our approach is able to avoid internal discontinuity and generate uniform pixel distribution. Moreover, the aspect ratio of our packing layout (i.e., 2 : 1 or 16 : 8) is very close to the majority standard aspect ratio 16 : 9. We expect our approach greatly improve 360 video quality and inspire more work along this direction.

In our current implementation, we do not perform any extra optimization. However, some optimization such as parallel computation, edge padding, etc., will potentially further speed up the entire performance.

Acknowledgments

This research work was supported by the NSFC (No.61631166002, No.61572196). We would like to thank Yao Zhao at the Intelligent Robot Motion & Vision Laboratory and Virtual Reality Innovation Laboratory of East China Normal University for his help with algorithm implementation. We also thank Wensong Li at VR8 Tech. Ltd.(Hangzhou, China) for helpful discussion.

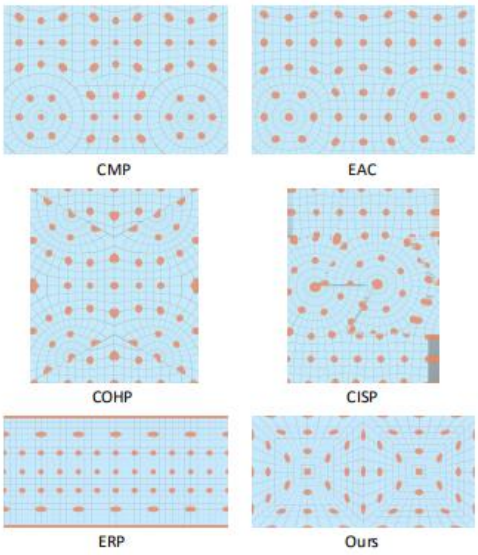

Figure 13: *Projections and packing layouts using various methods. Our method has better performance than the rest.*

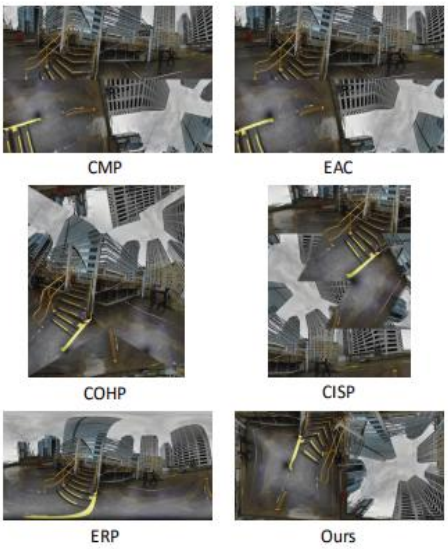

Figure 14: *360 Buildings: projections and packing layout using various methods.*

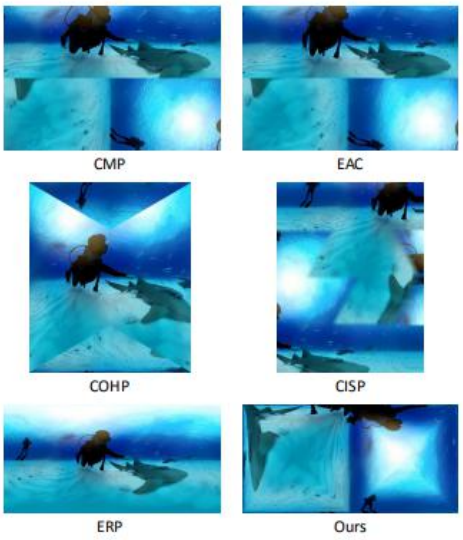

Figure 15: *360 Aquarium: projections and packing layout using various methods.*

Figure 16: *360 Indoor Conference: projections and packing layout using various methods.*

References

[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 Eurographics/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