

面向超多视点光场的虚拟立体内容生成技术综述

邢树军^{1,2}, 曹良才^{2*}, 桑新柱¹, 于迅博¹, 金国藩²¹北京邮电大学信息光子学与光通信国家重点实验室, 北京, 100876;²清华大学精密测试技术及仪器国家重点实验室, 北京, 100084

摘要 近年来,超多视点光场显示设备取得了快速发展,呈现出视点数目越来越多、分辨率越来越高和可视角度越来越大的特点,而能够呈现的虚拟场景也越来越复杂。对现有的超多视点光场显示设备的虚拟立体内容生成技术进行了梳理与总结,指出了各个技术适用的情景与优缺点。总体来说,面向超多视点的大规模虚拟场景的高质量实时光场内容生成技术还存在许多不足,需要研发全新的渲染体系才能突破现有算法的瓶颈。

关键词 图像处理; 光场显示; 超多视点渲染; 虚拟内容生成; 实时渲染

中图分类号 O436

文献标志码 A

doi: 10.3788/CJL202148.1509001

1 引言

2020 年,中国科幻大会展示了一款能够将三维虚拟物体动态悬浮于空中的超多视点光场显示设备^[1]。超多视点光场显示设备的商用化时代正在向我们走来,并展现出广阔诱人的应用前景。良好的光场显示设备,需要合适的虚拟内容生成来配合,才能展示出逼真动人的效果。

人眼之所以能看到世界中的物体,是因为人眼接收到了物体发出的光线,而光场就是三维世界中光线集合的完备表示。七维全光函数 $p(x, y, z, \theta, \varphi, \lambda, t)$ 可以完备地表述光场^[2]。斯坦福大学的 Levoy 等^[3]将七维全光函数简化为 (u, v, s, t) 4 维光场模型。虽然此模型并不能完备地表达三维空间中的所有光线,但是它剔除了与人眼前视方向相垂直方向的光线,使得数据量大大降低,又完备地表示了人眼成像的光线,此模型受到了业界的普遍认可。

光场技术主要分为光场采集技术和光场还原技术。光场采集技术又可分为实体相机采集技术和虚拟采集技术。虚拟采集是指利用计算机采集虚拟场景中的光场信息。针对光场虚拟采集技术得到的光场信息,可以根据超多视点光场显示设备的光学结

构,对光场信息进行编码或多角度投影,将得到立体显示的内容输出到相应的超多视点光场显示设备中进行显示。这种光场虚拟采集技术被称为超多视点光场虚拟采集技术。在光场还原技术中,无需佩戴眼镜、能够还原光场或部分还原光场的设备被称为超多视点光场显示设备。针对不同类型的超多视点光场显示设备和不同需求的场景,采用不同的虚拟立体内容生成方法。

本文将从超多视点光场显示设备的分类出发,详细介绍不同种类的超多视点光场显示设备在不同情景要求下的虚拟内容生成技术。除了对多视点图像进行渲染之外,本文还对光场编码以及投影阵列校正与同步进行了介绍,它们是超多视点光场内容生成技术的重要组成部分。

本文所采用的坐标系为 OpenGL 坐标系,即右手坐标系、列主序存储矩阵等。本文所指的大规模场景是指在视口(viewport)范围内的渲染面片数应不少于 200 万的虚拟场景;高质量的渲染标准至少应采用迪士尼原则的物理渲染标准。

2 超多视点光场显示设备的分类

超多视点光场显示设备主要包括光栅立体式超

收稿日期: 2021-03-15; 修回日期: 2021-04-21; 录用日期: 2021-05-24

基金项目: 国家自然科学基金(62075016, 2017YFB1002900)

通信作者: *clc@tsinghua.edu.cn

多视点光场显示设备、集成成像式超多视点光场显示设备和投影机阵列式超多视点光场显示设备三种,以及它们的悬浮展示形式。其视点数目一般不少于 50 个,单视点分辨率不应小于 640×480 。

2.1 光栅立体显示器

商用的光栅立体显示器^[4]可分为两种:一种是

基于狭缝光栅的立体显示设备,另一种是基于柱镜光栅的立体显示设备。前者由 LCD 面板或 LED 面板、透明垫板与狭缝光栅构成,后者由 LCD 面板或 LED 面板、透明垫板与柱镜光栅构成,后者的显示亮度明显高于前者。市场上绝大多数的光栅立体器是基于柱镜光栅分光原理生产制造的,如图 1 所示。

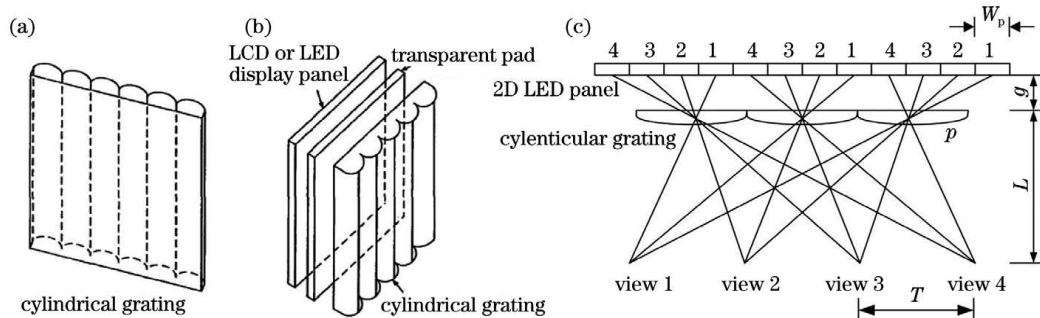


图 1 柱镜光栅立体显示器构造及其原理。(a)柱镜光栅;(b)柱镜光栅立体显示器;(c)柱镜光栅立体显示器分光原理
Fig. 1 Structure and principle of cylindrical grating stereoscopic display. (a) Cylindrical grating; (b) cylindrical grating stereoscopic display; (c) principle of light splitting for cylindrical grating stereoscopic display

通过柱镜光栅的分光作用,人的双眼可以看到不同的视点图像,从而在大脑中融合产生立体感。为了消除显示器像素的周期性与光栅周期性造成的摩尔纹,光栅一般被倾斜贴合于透明垫板上,其编码方法将在 4.1 节进行详述。

2018 年,北京邮电大学桑新柱团队发明了水平无跳变光栅光场显示装置,它由特种光源、液晶面板、狭缝光栅和全息功能屏组成,成功解决了自由立体显示设备跳变、串扰严重的问题^[5]。该装置的水平视角可达 100° ,远超现有的商用柱镜光栅立体显示器。文献[6]展示了分辨率为 22920×21600 的集群光场立体显示器,该显示器被认为是当前最高分辨率的光场显示设备,由一个由 15 台高性能渲染计算机构成的服务器驱动,每台计算机驱动一个 8K(7680×4320)光场立体单元,如图 2 所示。

目前显示效果较好的基于人眼跟踪的光栅立



图 2 集群光场立体显示器,它由 15 个 8K 液晶面板、15 个激光源和超大幅面柱镜光栅构成

Fig. 2 Cluster light-field stereoscopic display consists of 15 8K LCD panels, 15 laser sources, and super large-format cylindrical grating

体显示器为 Dimenco 公司和 Sony 公司生产的 8K 裸眼光栅立体显示器(图 3)。

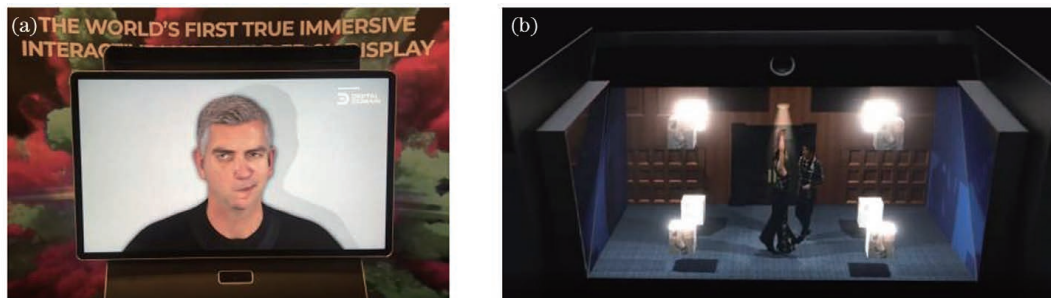


图 3 基于人眼跟踪的裸眼立体显示器。(a) Dimenco 8K 人眼跟踪裸眼立体显示器^[7];(b) Sony 人眼跟踪裸眼立体显示器^[8]
Fig. 3 Naked eye stereoscopic displays based on eye tracking. (a) Dimenco 8K eye-tracking naked eye stereoscopic display^[7]; (b) Sony eye-tracking naked eye stereoscopic display^[8]

2.2 集成成像光场显示设备

1908 年,布里埃尔·李普曼(Gabriel Lippmann)^[9]提出了集成成像的概念,之后他在 *Nature* 杂志上展示出了一张集成照片。集成成像是一种利用微透镜阵列来记录和再现三维图像的超多视点立体显示技术,该技术包括记录和再现两个过程。图 4(a)为采集过程,透镜阵列将三维场景记录在采集平面上,每一透镜对应着一个基元图像。图 4(b)则显示的是再现过程。首先将子图像阵列放置于显示平面上,使用相同的透镜阵列对场景进行再现。由于光路是可逆的,利用透镜阵列对空间中的光场进行还原,得到正确的立体像。其编码方案将在 4.2 节进行详述。

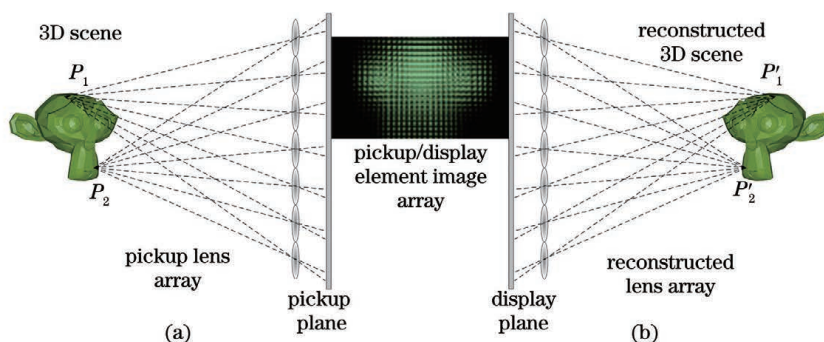


图 4 集成成像的采集与再现过程示意图。(a)采集过程;(b)再现过程

Fig. 4 Schematic diagram of acquisition and reproduction processes of integrated imaging. (a) Acquisition process; (b) reproduction process

近年来,集成成像与全息功能屏相结合的技术取得了不小的进展。如图 5(a)所示,2015 年广州泛彩溢公司发明了由 4K 商业 LCD 面板^[10]、透镜阵列和全息功能屏构成的集成成像光场显示设备,视点数目可达上万个。如图 5(b)所示,2020 年,这一设

备被改进为全视差光场电子沙盘^[11],其视点数目可达 144×144 个,总分辨率为 7680×8640 ,但是视点数目增多也带来了单视点还原图像分辨率的降低以及细节信息的丢失。

$$\frac{1}{f} = \frac{1}{g} + \frac{1}{l}, \quad (1)$$

式中: f 是透镜阵列的焦距; g 是显示平面与透镜阵列之间的距离; l 为显示平面上的像素经透镜折射后形成的中心深度平面与透镜阵列之间的距离。当 $g > f$ 时,设备工作在实模式下,还原出的立体图像为出屏的实像;当 $g < f$ 时,设备工作在虚模式下,还原出的立体图像为入屏的虚像;当 $g = f$ 时,设备工作在聚焦模式下,观察者观察到的立体效果既有入屏效果,又有出屏效果,但在这种模式下,分辨率会下降得非常严重。

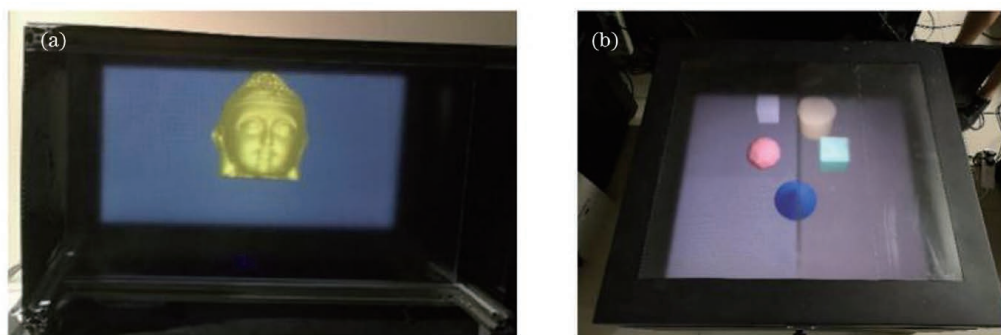


图 5 集成成像光场显示设备和全视差光场电子沙盘。(a)集成成像光场显示设备;(b)全视差光场电子沙盘

Fig. 5 Light field display of integral imaging, and full-parallax light field electronic sand table. (a) Light-field display of integral imaging; (b) full-parallax light field electronic sand table

2.3 投影机阵列式光场显示设备

投影机阵列式光场显示设备是将投影机阵列与定向扩散膜相结合的光场显示系统,一般需要大量

的投影机和多台计算机驱动,体积庞大,搭建困难,该系统需要同步机制的同步处理和投影校正算法的畸变处理,这一点会在 4.3 节详细描述。

Balogh 等^[12-13]研发了 HoloVizio 显示系统,它由 60 台分辨率为 1024×768 的投影机和平面向向散射膜(对角长度为 1.8 m)组成,系统水平视场范围为 50° ,角分辨率约为 0.8° ,如图 6(a)所示。

浙江大学刘旭团队构建了普通投影机阵列和圆



柱形定向散射屏相结合的三维显示系统^[14-16],环形排列投影机结合一个直径为 3 m 的弧形散射屏实现了视场范围为 360° 的大尺度三维光场显示,其中单个投影机分辨率为 640×480 ,角分辨率为 1° ,如图 6(b)所示。

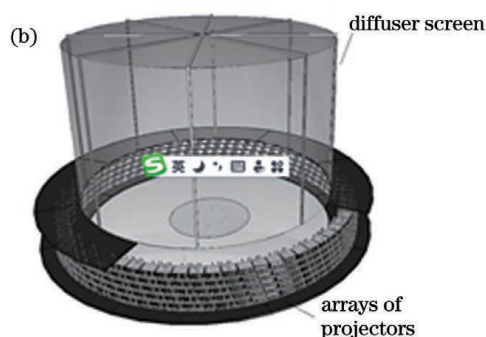


图 6 HoloVizio 显示系统和 360° 投影光场系统。(a) HoloVizio 显示系统;(b) 360° 投影光场系统




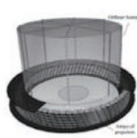
Fig. 6 HoloVizio display system and 360° projection light-field system. (a) HoloVizio display system; (b) 360° projection light-field system

表 1 列出了典型的几种较为成熟的超多视点光场显示设备。从表 1 可知,近年来超多视点光场显示设备呈现出视点数目越来越多、分辨率越来越高、可视角度越来越大的特点,而在现实应用中,需要显

示的虚拟场景也越来越复杂,所以需要更快速的高质量超多视点光场渲染算法以适应超多视点光场显示设备的发展。

表 1 典型的光场显示器参数比较

Table 1 Parameter comparison of typical light-field displays

Typical light-field display	Dimenco 8K naked eye tracking 3D display	Full-parallax electronic sand table	Cluster light-field stereoscopic display	360° projection light-field system
Picture				
Resolution	7680×4320	7680×8640	22920×21600	$720 \times 480 \times 36$
Viewpoint count	50-100	144×144	200	360
View angle / ($^\circ$)	50	115	80	360

2.4 光场立体显示的悬浮展示形式

光栅立体式、投影机阵列式、集成成像式光场显示设备可以与悬浮显示设备相结合,显示出悬浮的光场场景,此形式被称为光场立体显示的悬浮展示形式。现在比较实用的悬浮显示设备有离屏悬浮显示设备、负折射率板、逆反射悬浮成像设备。

离屏悬浮显示设备是由透镜、反射镜在空气中构建实像的设备。最简单的离屏空间悬浮显示设备^[17]为球面凹面反射镜,如图 7 所示,它会呈一个倒立的悬浮的实像,但是由于像差过大,悬浮出来的图像会发生严重畸变。

“海市蜃楼”模拟器由两个面对面的凹面反射镜

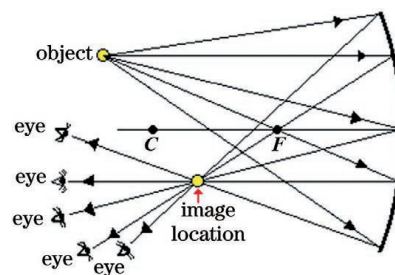


图 7 球面凹面反射镜悬浮^[18]

Fig. 7 Suspension of spherical concave mirror^[18]

组成,其面型均为抛物面^[18-20]。其中上面的凹面镜中间有一个开口,下面的凹面镜完全封闭。当把一个真实物体或显示器放在下端凹面镜抛物面焦点附

近,相应地,上端凹面镜的开口位置会出现该物体的实像,其原理如图 8 所示。



图 8 球面凹面反射镜悬浮

Fig. 8 Suspension of spherical concave mirror

“海市蜃楼”模拟器可以实现真实的漂浮影像,但是观看范围不大。只有在焦点处才是理想成像,在其他位置还原出来的悬浮图像存在严重畸变。

逆反射屏是可以将入射光线按照入射方向进行反射的一种屏,常应用于交通的警示牌。将图 9(a)中的凹面反射镜更换为逆反射屏^[21-22],以产生典型的逆反射悬浮显示的系统结构,如图 9(b)所示。

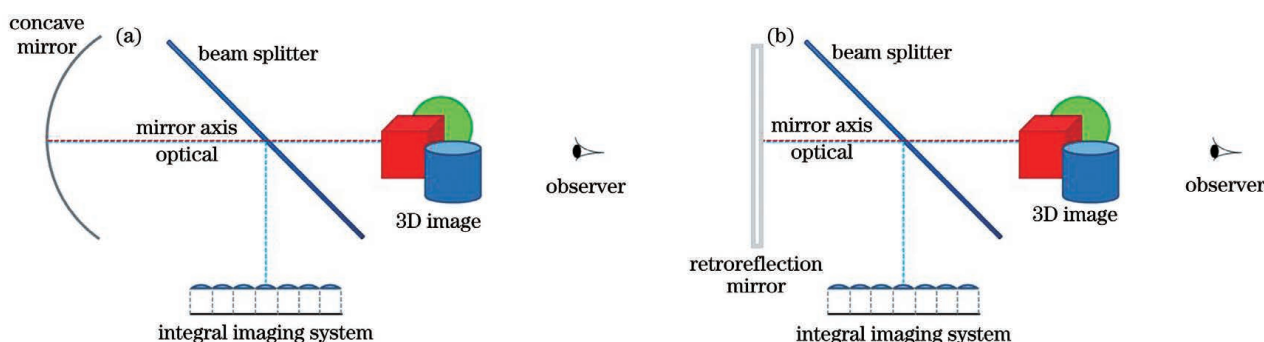


图 9 典型的基于逆反射的悬浮显示设备结构示意图^[23](柱镜光栅立体显示器可以替换为普通的显示器,则悬浮出的物体为一个二维平面)。(a)基于凹面镜的悬浮显示设备;(b)基于逆反射膜的悬浮显示设备

Fig. 9 Typical structural diagram of suspended display based on retroreflection^[23] (cylindrical grating stereoscopic display device can be replaced by general display, and suspended object is two-dimensional plane). (a) Levitating display device based on concave mirror; (b) suspension display device based on reverse reflection film

2020 年,悬浮光场设备被重新设计^[24],其原理与图 9(b)基本相同,在 70°可视范围内基本消除了像差,与其他光场立体显示器结合得到了人眼

可以接收的有视差的悬浮立体图像,如图 10 所示。其出屏距离可达 50 mm,像的尺寸为 30 cm×30 cm。

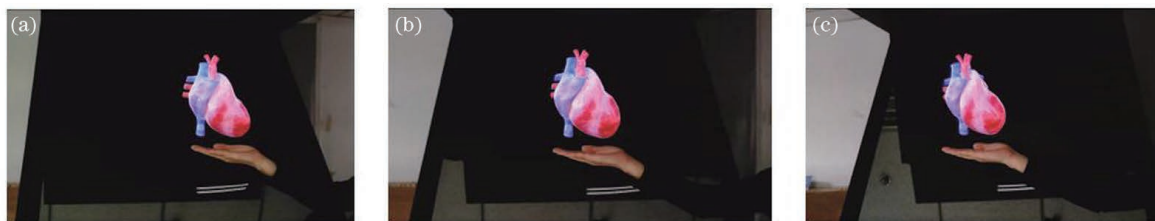


图 10 大尺寸高清晰度悬浮光场立体显示器。(a)左视点;(b)中视点图;(c)右视点图

Fig. 10 Suspended light field stereoscopic display with large size and high definition. (a) Left viewpoint image; (b) middle viewpoint image; (c) right viewpoint image

日本的 ASKA3D 公司^[25]和中国的东超公司^[26]独立地提出了一种具有负折射特性的玻璃板来实现空中成像。实现方式如图 11 所示,大量反射镜胶合在一起形成整齐排列的垂直镜面结构,虽然该结构看起来只是简单的两层式镜面结构,但是要

实现几百微米的规则性图案并不容易,由于制作工

艺复杂、批量化难度大,其成本也很难降低。总体来说,目前光栅立体式、投影机阵列式超多视点光场和它们的悬浮展示形式的显示效果较好,能够动态展示虚拟场景,能在不久的将来进入实用阶段,而集成成像光场显示设备的单视点分辨率还有待提高。

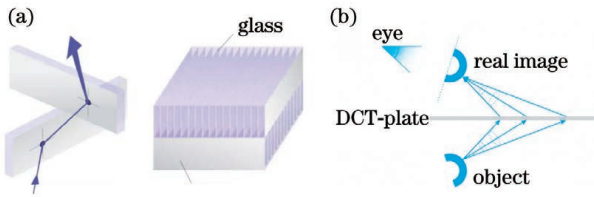


图 11 具有负折射特性的玻璃板的基本原理。(a)结构图；(b)光路图

Fig. 11 Basic principle of glass panel with negative refraction. (a) Structural chart; (b) light-path diagram

3 超多视点虚拟内容的生成方法

采集空间中的光场信息最简单的方式为在虚拟空间中设置一个摄影机阵列，每一个摄影机对应一个视点，利用虚拟摄影机生成视点图像。在介绍超多视点虚拟内容的生成算法之前，本文先对虚拟摄影机的排列方式和参数设置进行简单介绍。

3.1 虚拟摄影机的排列方式

图 12 说明了四维光场的一种采集过程，通过从 st 平面向 uv 平面上一点投射光线，生成相应的视

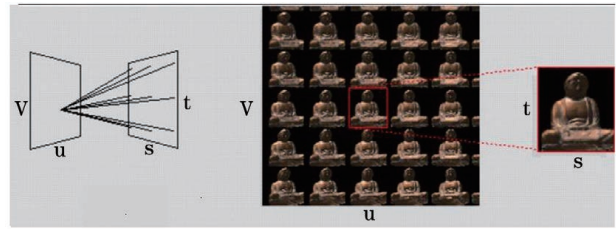


图 12 四维光场采集示意图^[3]

Fig. 12 Schematic diagram of four-dimensional light field acquisition^[3]

点图像，假设 uv 平面上有 $M \times N$ 个投影灭点，可以将 $M \times N$ 虚拟摄影机放置在灭点上，根据投影关系设置相机参数。每一个虚拟摄影机可以生成一个视点图像，在计算机图形学中，用模型视图矩阵去设置摄影机的位姿(即外参)，用投影矩阵设置视野范围、远近裁切面(即内参)。投影矩阵还可以将相机坐标系转化为归一化设备坐标系，以利于后期的裁切操作。

对于空间中一点 P ，其在世界坐标系的坐标为 $(p_{wx}, p_{wy}, p_{wz}, 1)$ ，那么它在相机坐标系下的坐标为 $(p_{vx}, p_{vy}, p_{vz}, 1)$ ，两者存在以下关系：

$$(p_{vx}, p_{vy}, p_{vz}, 1) = \begin{pmatrix} u_x & v_x & n_x & 0 \\ u_y & v_y & n_y & 0 \\ u_z & v_z & n_z & 0 \\ e_x & e_y & e_z & 1 \end{pmatrix} (p_{wx}, p_{wy}, p_{wz}, 1), \quad (2)$$

式中： (e_x, e_y, e_z) 是相机在世界坐标中的位置，相机坐标系的 u, v, n 轴分别对应于世界坐标系的 x, y, z 轴，如图 13 所示。模型视图矩阵可表示为

$$M_v = \begin{pmatrix} u_x & v_x & n_x & 0 \\ u_y & v_y & n_y & 0 \\ u_z & v_z & n_z & 0 \\ e_x & e_y & e_z & 1 \end{pmatrix}. \quad (3)$$

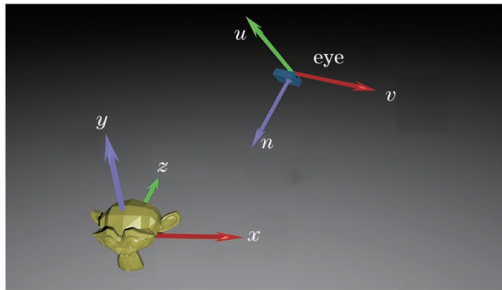


图 13 相机坐标系与世界坐标系的转换

Fig. 13 Transformation from camera coordinate system to world coordinate system

相机视景体的 6 个参数 n, f, l, r, b, t ，分别代表远、近裁切面到相机的距离，以及近切面中的左、右、上、下边到相机朝向矢量的垂直距离。只要相机视景体确定了，其投影矩阵也就可以确定了。从图 14 可以发现其投影平面在近切面上，这与反透视

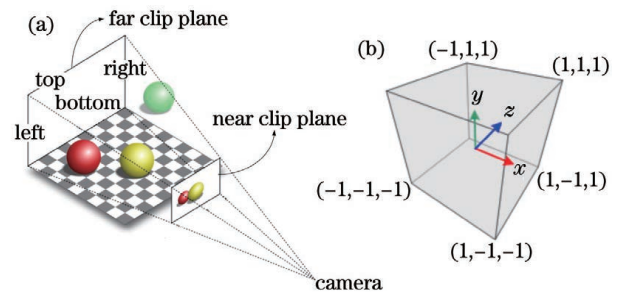


图 14 线性投影变换和规范化设备坐标系。(a)线性投影变换；(b)规范化设备坐标系

Fig. 14 Linear projection transformation, and normalized device coordinate system. (a) Linear projection transformation; (b) normalized device coordinate system

投影有明显的不同。

在中心相机中, $r = -l, b = -t$, 其透视投影矩阵为

$$\mathbf{M}_{pc} = \begin{pmatrix} n/r & 0 & 0 & 0 \\ 0 & n/t & 0 & 0 \\ 0 & 0 & \frac{-(f+n)}{f-n} & -1 \\ 0 & 0 & \frac{-2fn}{f-n} & 0 \end{pmatrix} \quad (4)$$

如图 15 所示, 虚拟摄影机阵列的视景体设置有

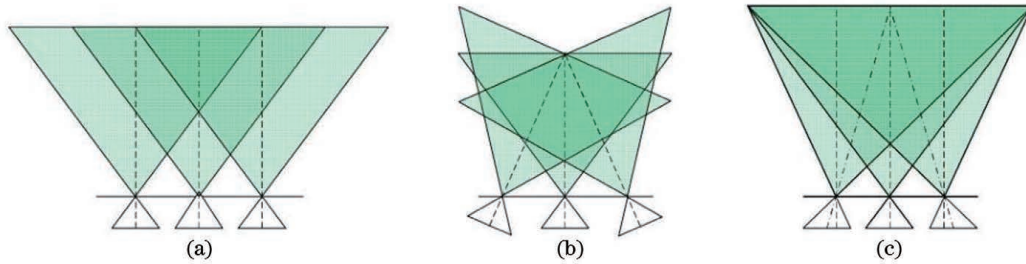


图 15 多视点立体摄影机视景体的设置方法。(a) 平行式; (b) 汇聚式; (c) 错切式

Fig. 15 Setting methods of view frustum for multi-view stereo camera. (a) Parallel type; (b) cluster type; (c) sheared type

上述三种虚拟摄影机的设置可应用于光栅式、投影式光场显示设备和以矩形排列的透镜阵列集成成像式光场显示设备, 并且这三种排布方式既可以面向水平视差光场, 也可以面向全视差光场。而对于蜂窝排布或不规则排布的集成成像光场显示设备, 一般不采用虚拟摄影机阵列进行采集, 而是利用光线跟踪的方法去解决, 这一点将在 3.4 节进行详细介绍。

假设使用 $M \times N$ 个相机阵列去采集 4 维光场信息, 相机阵列到零视差平面的距离为 d_{scr} 。已知中心相机的模型视图矩阵为 \mathbf{M}_{mc} , 透视投影矩阵为 \mathbf{M}_{pc} 。

若要进行平行式采集, 只需要将错切投影矩阵 \mathbf{M}_{pmm} 设置为 \mathbf{M}_{pc} 即可。对于汇聚式采集来说, 相机 (m, n) 的 \mathbf{M}_{pmm} 等于中心相机透视投影矩阵 \mathbf{M}_{pc} , 只需使用各大引擎提供的 LookAt 函数去计算模型视图矩阵就可以了。输入相机的位置, 汇聚点位置和相同的向上向量, 就可以生成模型视图矩阵 \mathbf{M}_{mn} :

$$\mathbf{M}_{mn} = \mathbf{M}_{mc} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ d_{mx} & d_{ny} & 0 & 1 \end{pmatrix} \quad (5)$$

对于错切式采集来说, 在相机阵列中第 (m, n) 个相机的模型视图矩阵和错切投影矩阵^[30]需分别

以下三种方法: 平行式设置^[27]、汇聚式设置^[28]和错切式设置^[29]。对于平行式视景体设置来说, 虚拟摄影机的光轴相互平行, 但它只能显示出屏立体效果, 而且公共区域狭小。而对于汇聚式设置方式, 只有屏幕的中心线位置没有垂直视差, 两边的垂直视差最大。人眼感到两边的立体图像向屏幕内凹陷, 造成立体的畸变。错切式设置方法既能显示出屏立体显示效果, 又能显示入屏立体显示效果, 并且没有汇聚式设置产生的畸变。平行式视景体与错切式视景体之间存在三维错切关系。

设置为

$$\mathbf{M}_{mn} = \mathbf{M}_{mc} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ d_{mx} & d_{ny} & 0 & 1 \end{pmatrix}, \quad (6)$$

$$\mathbf{M}_{pmm} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ d_{mx}/d_{scr} & d_{ny}/d_{scr} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{M}_{pc}, \quad (7)$$

式中: d_{mx} 、 d_{ny} 分别代表相机阵列中第 (m, n) 个相机到中心相机的水平距离和垂直距离。

3.2 逐视点渲染法

逐视点渲染方法是指在虚拟空间中以一定规则摆放虚拟摄影机阵列, 并逐个渲染出视点图像, 最终合成光场编码图像, 或者将每一个视点图像输出至相应的投影机上, 多角度投影在相应的光场屏幕上。本方法既可以采用栅格化的方法, 也可以采用光线跟踪的方法。

3.2.1 串行逐视点渲染法

前期的逐视点渲染方案如图 16 所示, 需要以串行的方式执行。本方法的中间过程中生成的每个视点图像存在大量冗余, 生成时间随着视点数目的增多而线性增加。该方法一般应用在 100 个视点以内、单视点图像分辨率小于 640×480 、场景面片数

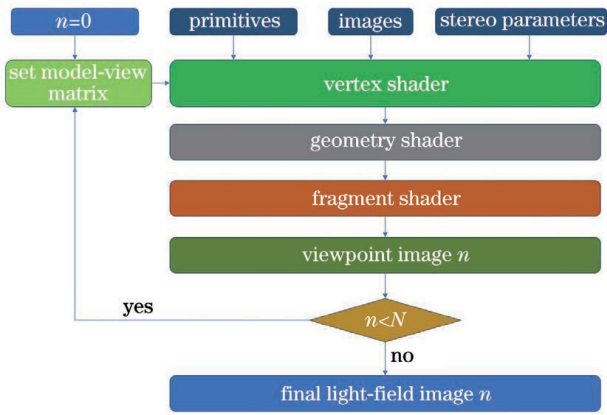


图 16 串行栅格化逐视点渲染法流程图

Fig. 16 Flow chart of view-by-view rendering based on serial rasterization

小于 10^7 的光场实时渲染。光线跟踪渲染一般不使用逐视点的渲染方法,因为光线跟踪渲染相对于栅格化渲染十分缓慢,当渲染成千上万个视点,一帧 8K 的光场图像可能要花费几天的时间。

3.2.2 基于几何着色器的逐视点渲染法

文献[31]中引入了采用几何着色器复制图元的方法,以六视点图像的渲染为例,如图 17 所示。

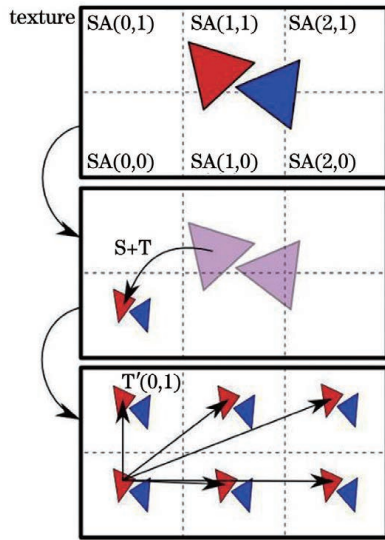


图 17 基于几何着色器的图元复制方法^[31]

Fig. 17 Copy method of primitives based on geometric shaders^[31]

基于几何着色器的逐视点渲染法可分为三个步骤进行:1)在几何着色器中根据视点数目对基本的图元进行复制;2)根据视点参数对复制后的图元进行变换操作;3)根据相应的摄影机参数进行裁切操作,可一次性地在纹理上渲染出多个视点的图像。该方法可以节省多视点渲染时间,但缺点也很突出,即基本图元的复制会导致显存的占用率大大

增加,这是以空间换时间的做法,只能用于显示小场景,大规模场景很可能因图元数目过多导致显存溢出,程序崩溃。

3.2.3 基于 GPU Instancing 的逐视点渲染法

GPU Instancing 是一种用来提高渲染大量物体效率的技术^[32-33],随着人们对三维场景品质需求的提升,需要在场景里绘制越来越多的物体,这里面主要涉及两个方面的性能瓶颈:一是 CPU 对 GPU 提交数据的次数(包括设置数据缓冲区,渲染状态以及调用对渲染原语的绘制 drawcall);二是在 GPU 上的绘制(主要包括顶点处理和像素绘制)。随着场景物体的增多,CPU 和 GPU 的压力都会上升。

如图 18 所示,在场景中有大量重复物体需要绘制时,使用 GPU Instancing 技术只需要设置一次原始数据缓冲区,调用一次 drawcall 即可绘制出物体。与基于几何着色器的逐视点渲染算法相比,该技术的显存需求不会随着视点数目的增加而增加。所以,本算法相对于几何着色器算法能够渲染更大规模的场景,但是已经对场景使用了 GPU Instancing 技术绘制重复虚拟物体,就不能够再使用它进行多视点绘制,因此无法完成光场内容的生成。

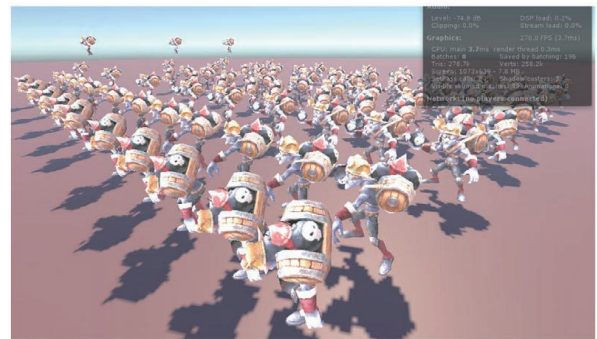


图 18 Unity 3D 中的 GPU instancing 技术,可解决大量重复物体的计算效能问题

Fig. 18 GPU instancing technology in Unity 3D, solving computational efficiency problem of generous repetitive objects

3.3 反透视逐透镜渲染法

图 19(a)展示是普通的透视投影,将物体向靠近投影灭点的裁切面进行投影,物体朝向灭点的面会被记录下来。而图 19(b)展示的是反透视投影,将物体向远离投影灭点的裁切面进行投影,背向灭点的面会被记录下来,得到图像呈现出近小远大的视觉特性,此为反透视投影的基本原理。反透视投影矩阵为

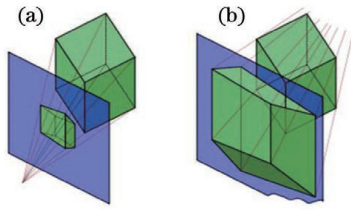


图 19 透视投影的原理。(a)普通的透视投影；
(b)反透视投影

Fig. 19 Principles of perspective projections. (a) Common perspective projection; (b) reverse perspective projection

$$M_{pc} = \begin{pmatrix} n/r & 0 & 0 & 0 \\ 0 & n/t & 0 & 0 \\ 0 & 0 & \frac{f+n}{f-n} & -1 \\ 0 & 0 & \frac{2fn}{f-n} & 0 \end{pmatrix}. \quad (8)$$

反透视特性的应用十分少见^[34],只存在于宗教画和特殊建筑的效果图中^[35],如图 20 所示。



图 20 反透视绘制应用。(a)宗教画;(b)特殊建筑

Fig. 20 Applications of inverse perspective drawing. (a) Religious painting; (b) special construction

逐视点渲染法需要先生成各个视点的图像,再进行编码输出,中间生成的视点图像存在大量冗余,对于集成成像系统来说,尤其如此。工作在实模式下的集成成像设备,其基元图像为反透视像,可以将每一个透镜当作一个反透视摄影机,逐透镜地进行渲染,得到对应的基元图像,最终拼接合成一整幅的实模式图像,将其直接加载到集成成像的 LED 或 LCD 面板上进行显示。

反透视投影基本的绘制步骤为两步。第一步为阴面消隐^[36],普通的线性透视投影将背向摄影机的一面定义为阴面进行消隐,将反透视朝向摄影机的一面定义为阴面进行消隐,如图 21 所示。第二步为反画家算法^[37],与普通的线性透视投影的画家算法不同,反画家算法首先绘制距离较近的场景,然后用绘制距离较远的场景覆盖较近的部分。在运用反画家算法之前,需要对深度进行排序。

反透视逐透镜渲染法只适用于工作在实模式下的集成成像光场设备,虽然消除了信息冗余,但随着透镜数据的增多,反透视渲染需要重复执行,将消耗大量计算资源。它仍然需要与几何着色器算法或 GPU instancing 技术相结合达到加速的目的,在此过程中也就不可避免地引入了其算法或技术带来的固有缺陷。

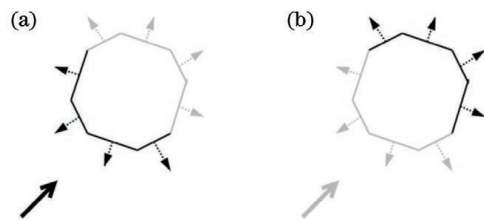


图 21 普通线性透视投影的阴面消隐和反透视投影的阴面消隐。(a)普通线性透视投影的阴面消隐；
(b)反透视投影的阴面消隐

Fig. 21 Shadow elimination of general linear perspective projection, and shadow elimination of inverse perspective projection. (a) Shadow elimination of general linear perspective projection; (b) shadow elimination of inverse perspective projection

3.4 光线跟踪法

光线跟踪算法是另一种无冗余的光场内容生成算法,它适用于任意类型的光场显示设备,可以生成质量很高的光场图像,将其直接加载至光场的二维显示面板或投影机即可显示,但是生成速度较慢。虽然 2018 年英伟达在显卡中引入了实时光线跟踪模块,推出了实时光线跟踪图形接口 Optix,但是仍然不能使用本方法实现虚拟大场景光场的实时渲染。

光线跟踪算法一般由三部分构成^[38],即光线的生成、光线的碰撞和像素的着色。传统显示设备的光线跟踪与光场显示设备的光线跟踪的区别在于光线的生成部分。如图 22 所示,光线从虚拟摄影机出发经透镜到达基元图像像素,像素和光线满足一一对应的关系。通过光场显示器的像素生成光线,并进行碰撞检测,最后通过着色程序就可以生成相应的光场图像。

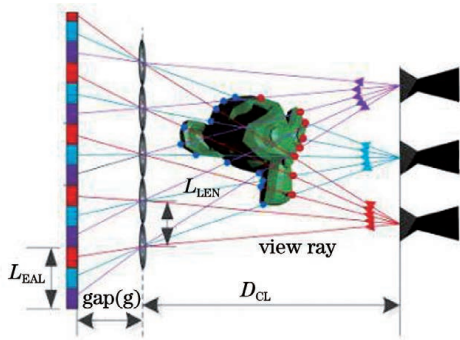


图 22 基于光线跟踪的集成成像算法^[39]

Fig. 22 Integrated imaging algorithm based on ray tracing^[39]

光线跟踪像素与光线映射的任意性可以用来修正装配误差^[40],提高光场显示的质量,也可以适配各种不规则的光场显示设备。光线跟踪具有天然的并行性,可以很方便地提高光线跟踪的效率。但光线与复杂场景的碰撞检测是一种高度消耗计算资源的操作,难于并行化。光线跟踪的计算复杂度与屏幕的分辨率大小呈正相关,现有的实时光线跟踪硬件管线只能面向 2K 以下的光场显示设备且可绘制的场景非常简单^[41]。在进行体数据渲染时,光线跟踪算法是一种常用的算法,文献[42-43]中详细讲述了基于光线跟踪的体数据超多视点渲染,该方法与面数据的渲染大同小异。光栅式光场显示器是以子像素编码的,在总分辨率一定的情况下,其计算资源的消耗量是集成成像光场渲染计算资源消耗量的 3 倍^[44]。

近些年,光线跟踪技术还与人工智能(AI)超分辨率技术相结合^[45],用于快速生成高质量光场图像,达到了可交互的效果。在虚拟物体运动过程中光线采样率低,静止时光线采样率高,此方法离真正的高质量实时内容生成还有很大差距。

3.5 基于深度信息的超多视点光场渲染

基于深度信息的渲染(DIBR)在虚拟场景的渲染过程中十分常用^[46],不适用于工作在实模式下的集成成像光场显示设备。DIBR 算法是利用深

度信息和其他附加信息、通过插值产生其他视点的图像,它有效地降低了图形渲染的复杂度,渲染速度大大加快,但是也会造成渲染质量的下降。

根据参考视点的数目,DIBR 可分为两类:一类为单参考视点的 DIBR^[47-49],另一类为多参考视点的 DIBR^[50]。单参考视点的 DIBR 可以只使用一幅深度参考图像和彩色图像生成光场所需要全部视点,但当视角较大时,空洞较大,填补困难^[51],它只适用于 10°以内观看视角的光场显示设备。多参考视点的 DIBR 需要多个深度参考视点,能够有效地增大视角,消除空洞^[52]。多参考视点一般使用左、右两个视点进行插值,得到中间视点。DIBR 技术具有带宽需求小、输入图像数量少和绘制速度快的优点。单参考视点 DIBR 技术的映射速度快,双参考视点的 DIBR 技术能够利用左、右视图实现对遮挡区域的信息互补,如图 23 所示。无论单参考视点的 DIBR 技术还是双参考视点的 DIBR 技术都仍然存在空洞、重采样、重叠和伪影问题。

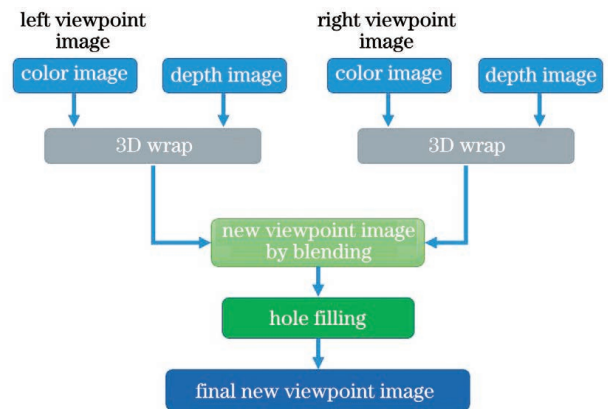


图 23 双边 DIBR 算法的流程

Fig. 23 Flow chart of bilateral DIBR algorithm

此外,基于深度的计算方法会产生错误的光照信息,尤其是在对各向异性材质物体进行渲染时,产生的光照信息和真实空间中的光照特性相差较大。为了解决这一问题,2020 年,一种基于延迟着色的超多视点光场图像生成方法被提出^[53],其参考图像由一对或多对彩色图像、深度图像、高光图像和环境图像构成。虽然该方法解决了光照问题,但其和基于深度图像的绘制一样仍然会导致多视点画质的下降。

3.6 基于几何相关性的超多视点光场渲染

上述超多视点裸眼光场渲染算法均没有利用视点间的几何相关性来达到去除信息冗余的目的。图 24 是一组错切多视点图像和它对应的横

切面图像,即 Epipolar Plane Image (EPI) 图像^[54],它反映出虚拟物体在多视点图像之间的几何相关性。

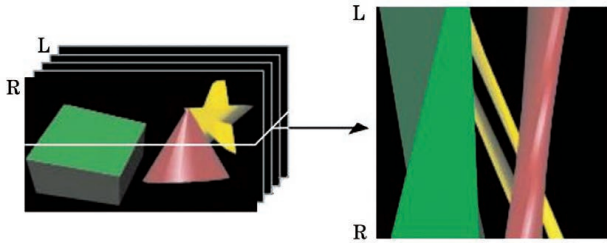


图 24 多视点图像与其对应的 EPI 图像^[55]

Fig. 24 Multi-view images and corresponding EPI image^[55]

假设三个点场景组成三维场景,虚拟摄影机阵列为错切式排列,如图 25 所示,其对应的 EPI 图像为三条直线。

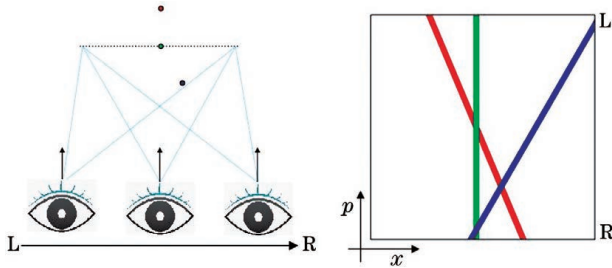


图 25 虚拟点在错切排列下的光场采集以及对应的 EPI 图像

Fig. 25 Light field acquisition of virtual points with staggered arrangement and corresponding EPI image

当点位于零平面时,所对应的 EPI 图像为竖直的直线;若点位于零平面内,则对应的 EPI 图像是斜率为正的直线;若点位于零平面外,则对应的 EPI 图像是斜率为负的直线。因此,点的渲染就可以转化为最左侧虚拟相机和最右侧相机对这一点的渲染,并在 EPI 图像上由这两视点生成相应的 EPI 直线,最终再转换为视点图像,这样就会大大简化渲染的流程。这是基于几何相关性的超多视点光场渲染的基本原理。

2020 年,一款利用几何相关性原理并针对多边形栅格化的超多视点渲染管线被提出^[56],在场景复杂度为 200 万面片、总分辨率为 7680×4320 的情况下,渲染速率达到 60 frame/s,但是其光照模型和纹理映射等诸多因素没有被考虑,不能和传统的图形硬件渲染管线相兼容,它表现出惊人的效率提升效果,后面需要重点探索光场材质、光照的几何相关性原理和与之相适应的图形渲染硬件结构,以保证光场图像的高质量生成。它也可以和反透视相

结合,快速生成用于工作在实模式下的集成成像光场设备的内容。本方法可适用于其他任意类型的光场显示设备。

4 编码、校正与同步

光栅式立体显示设备和集成成像显示设备需要通过编码将得到的多视点图像生成最终的光场立体图像,并将其加载到 LED 或 LCD 面板上进行显示。而对于投影式光栅立体设备来说,需要将各个视点的图像加载到每个投影机上,由于视点数目较多,需要相应的同步机制进行同步;此外,也需要对投影过程中存在的投影畸变进行畸变校正。

编码是指将视点图像合成最终光场图像的过程。对于光线跟踪生成算法来说,编码与光线跟踪过程同步进行,根据编码参数生成相应的光线,直接生成光场图像。

4.1 光栅式光场显示设备的编码算法

光栅式立体显示设备的参数如图 26 所示,其中,

$$N_{kl} = \frac{(k - k_{off} + V_{shift} - 3l \tan \alpha) \bmod D_x}{D_x} N_{tot} \quad [57], \quad (9)$$

式中: (k, l) 为子像素索引; α 为倾角; D_x 为水平狭缝周期内的视点数; N_{tot} 为总视点数; k_{off} 为子像素左上角点到 $(0, 0)$ 点的水平距离。在已知子像素索引 (k, l) 的情况下,可以通过倾角 α 、水平狭缝周期内视点数 D_x 、总视点数 N_{tot} 、以及 k_{off} 计算得到相应的视点数 N_{kl} ,这里的 V_{shift} 值可以应用于人眼跟踪的立体显示^[58]。为了改善显示效果,文献^[59]引入了加权编码操作,即在一个子像素内根据不同的比例混合多个视点的信息。

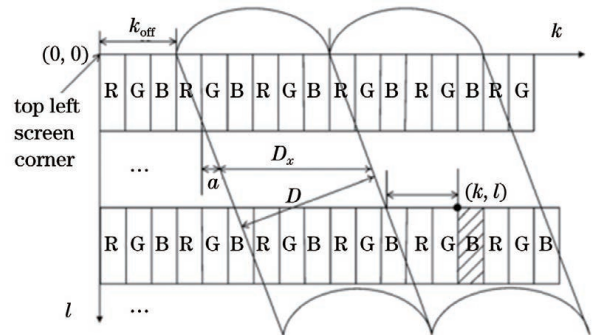


图 26 光栅式立体显示器参数示意图

Fig. 26 Parameters of grating stereoscopic display

4.2 集成成像光场显示设备的内容编码算法

在文献^[40]中,已知集成成像中的每一个透镜的位置为 p_{ij} ,其中 i, j 对应的是透镜编号。计算

LCD 面板或 LED 面板中每一个像素对应距离最近的透镜,这个透镜对应的像素区域称为基元图像。基元图像的形状可以为正方形、正六边形或者不规则图形。计算基元图像的中心位置,以及其轴向外接矩形的长和宽,如图 27 所示。根据对应透镜的参数生成相应长宽的图像,经过中心配准、裁切后得到基元图像,最后由基元图像拼接成集成成像光场图像。

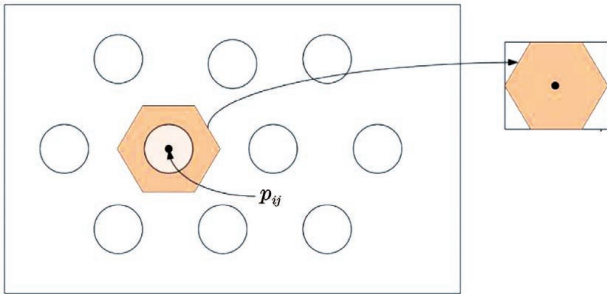


图 27 光场基元图像的生成

Fig. 27 Generation of elemental images in light field

4.3 投影式光场的校正与同步算法

在对投影式光场进行显示前需要对投影机的投影区域进行校正,如图 28 所示,投影机 B 的原始图像的投影区域为 $A'B'C'D'$,其顶点坐标用 (X, Y) 表示;其目标投影区域为 $ABCD$,其顶点坐标用 (x, y) 表示。

在进行校正时,常用到单应性公式^[60]:

$$(x, y) = \left(\frac{h_1 X + h_2 Y + h_3}{h_7 X + h_8 Y + 1}, \frac{h_4 X + h_5 Y + h_6}{h_7 X + h_8 Y + 1} \right), \quad (10)$$

将 4 对顶点对应的数据代入(10)式,联立方程组,求出单应性参数 $(h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8)$ 。

通过输入原始图像的坐标,就可以计算出对应目标图像的坐标,这一步骤很容易在渲染管线中的片元着色器中进行操作。

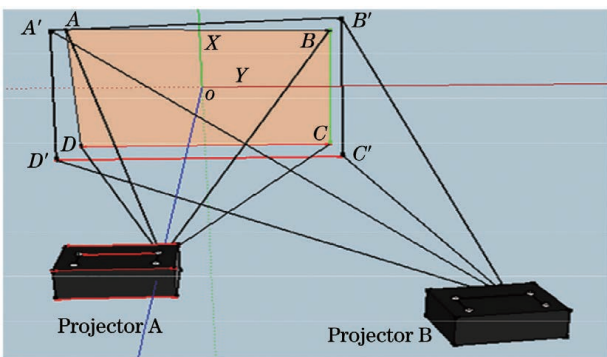


图 28 投影机的校正^[61]

Fig. 28 Correction of projectors^[61]

由于投影阵列光场显示设备是用集群计算机驱动的,一般采用星型连接的方式,同一局域网中存在一个主机和众多从机。主机通过 UDP BroadCast 的方式^[62]向从机广播中心摄影机的模型视图矩阵和场景中动态虚拟物体的变换矩阵。从机接收到相应的矩阵后,配置其对应相机的模型视图矩阵、投影矩阵和动态虚拟物体的变换矩阵^[63]。由于场景矩阵的传输量极小,延迟可控制在 10 ms 以内。

5 结束语

通过上面的介绍,本文将多种超多视点光场的渲染算法、编码、校正与同步算法在各种光场显示设备上运行的特性进行总结,如表 2 所示。

表 2 中的 A、B、C、D 与表 1 中的光场显示设备相对应,分别代表 Dimenco 8K 人眼跟踪裸眼立体显示器、全视差光场电子沙盘、集群光场立体显示器、浙江大学 360°投影光场系统。表 2 中的“!”代表算法无法完成对应设备虚拟内容的生成,而不参与速度或质量排名;“×”代表不需要本算法或没有相应特性;“√”代表本算法拥有相应特性;“*”代表视角无限制;“↓”代表图形质量有损失,而“↓↓”代表图形质量损失严重,“-”代表图像质量无损失。 α, β, γ 分别代表光栅式光场显示设备的编码算法、集成成像光场显示设备的内容编码算法,以及投影式光场的校正与同步算法。

从表 2 对比可知,现阶段渲染质量最高的算法为光线跟踪算法(算法 2 和 4),非实时的应用可使用这一算法;现阶段速度最快的渲染算法是基于几何相关性的渲染算法(算法 10),但是它与现有的硬件渲染管线不兼容,渲染过程中光照、遮挡都需要做特殊处理;目前最为实用的两种实时渲染技术为基于延迟着色的渲染(算法 8)和基于 GPU instancing 的光场渲染(算法 9),但是前者出现质量图像的下降且不能展示透明虚拟物体,后者无法实现大规模场景的高质量渲染。总体来说,面向超多视点的大规模虚拟场景的实时光场内容生成技术还尚未完全成熟,需要探索全新的渲染体系才能突破现有算法的瓶颈。基于几何相关性的渲染算法表现出惊人的效率提升效果,是未来光场渲染算法研究的主流方向,因此需要重点探索光场材质,光照的几何相关性原理,和与之相适应的图形渲染硬件结构,以保证超多视点光场图像的实时高质量生成。

参 考 文 献

- [1] Sang X. CCTV interview on China science fiction conference 2020 [EB/OL]. (2020-10-31) [2020-03-14]. http://file.cgt3d.cn:8013/cctv_interview_bupt.mp4.
- [2] Moon P, Spencer D E. Theory of the photic field[J]. Journal of the Franklin Institute, 1953, 255(1): 33-50.
- [3] Levoy M, Hanrahan P. Light field rendering [J]. ACM, 1996: 31-42.
- [4] Wang A H, Wang Q H. Overview on autostereoscopic display device based on parallax barrier or lenticular lens [J]. Advanced Display, 2009, 20(10): 12-17.
- [5] Wang P R, Sang X Z, Yu X B, et al. Demonstration of a low-crosstalk super multi-view light field display with natural depth cues and smooth motion parallax [J]. Optics Express, 2019, 27(23): 34442-34453.
- [6] Ultra high resolution autostereoscopic display [EB/OL]. (2020-08-20) [2021-03-14]. http://file.cgt3d.cn:8013/uhr_8k.mp4.
- [7] Dimenco company, eye tracking 8k autostereoscopic display [EB/OL]. (2020-07-12) [2021-03-14]. <http://file.cgt3d.cn:8013/Dimenco8K.mp4>.
- [8] Sony company, eye-sensing light field display [EB/OL]. (2020-07-12) [2021-03-14]. <http://file.cgt3d.cn:8013/Sony8K.mp4>.
- [9] Lippmann G. La photographie integrale[J]. Comptes Rendus de l'Académie des Sciences, 1908, 146, 446-451.
- [10] Fan F C, Choi S, Jiang C C. Demonstration of perfect holographic display on commercial 4K plane displayer [C] // Digital Holography and Three-Dimensional Imaging 2015, May 24-28, 2015, Shanghai, China. Washington, D. C.: OSA, 2015: DW3A.4.
- [11] Sang X Z, Gao X, Yu X B, et al. Interactive floating full-parallax digital three-dimensional light-field display based on wavefront recomposing [J]. Optics Express, 2018, 26(7): 8883-8889.
- [12] Balogh T, Zanetti G, Bouvier E, et al. An interactive multi-user holographic environment [C] // ACM SIGGRAPH 2006 Emerging technologies on-SIGGRAPH '06, July 30-August 3, 2006, Boston, Massachusetts. New York: ACM Press, 2006: 18-42.
- [13] Agocs T, Balogh T, Forgacs T, et al. A large scale interactive holographic display [C] // IEEE Virtual Reality Conference (VR 2006), March 25-29, 2006, Alexandria, VA, USA. New York: IEEE Press, 2006: 311.
- [14] Zhong Q, Chen B S, Li H F, et al. Multi-projector-type immersive light field display [J]. Chinese Optics Letters, 2014, 12(6): 060009.
- [15] Zhong Q, Peng Y F, Li H F, et al. Multiview and light-field reconstruction algorithms for 360° multiple-projector-type 3D display [J]. Applied Optics, 2013, 52(19): 4419-4425.
- [16] Peng Y F, Li H F, Zhong Q, et al. Large-sized light field three-dimensional display using multi-projectors and directional diffuser [J]. Optical Engineering, 2013, 52(1): 017402.
- [17] McKay S, Mason S, Mair L S, et al. Membrane-mirror-based display for viewing 2D and 3D images [J]. Proceedings of SPIE, 1999, 3634: 144-155.
- [18] Ray Diagrams-Concave Mirrors. Reflection and the ray model of light-lesson 3-concave mirrors [EB/OL]. [2020-03-14]. <https://www.physicsclassroom.com/class/refln/Lesson-3/Ray-Diagrams-Concave-Mirrors>.
- [19] Silva O H M, Laburú C E. Uma alternativa de baixo custo ao experimento de óptica denominado comercialmente "magic hologram-mirage 3D" [J]. Caderno Brasileiro De Ensino De Física, 2014, 32 (1): 246.
- [20] Mahfud J, Matsumaru T. Interactive aerial projection of 3D hologram object [C] // 2016 IEEE International Conference on Robotics and Biomimetics (ROBIO), December 3-7, 2016, Qingdao, China. New York: IEEE Press, 2016: 1930-1935.
- [21] Tsuchiya K, Koizumi N. Aerial imaging by retro-reflection for mid-air image display and camera viewpoint transfer [J]. IEEE Access, 2021, 9: 25110-25117.
- [22] Kim Y M, Jung K M, Min S W. Analysis of off-axis integral floating system using concave mirror [J]. Journal of the Optical Society of Korea, 2012, 16 (3): 270-276.
- [23] Yamamoto H, Tomiyama Y, Suyama S. Floating aerial LED signage based on aerial imaging by retro-reflection (AIRR) [J]. Optics Express, 2014, 22 (22): 26919-26924.
- [24] Du J Y, Sang X Z, Yu X B, et al. Large viewing angle floating three-dimensional light field display based on the spatial data reconstruction (SDR) algorithm [J]. Optics Communications, 2020, 475: 126229.
- [25] A 3D Company. Aska3D plate [EB/OL]. [2020-03-14]. <https://aska3d.com/zh-cn/products.php>.
- [26] Huang Z G, Fan C, Han D C, et al. A kind of flat plate optical structure: CN207502815U [P]. 2017-07-26.

- 黄志刚, 范超, 韩东成, 等. 一种实现空气成像的光学平板结构: CN207502815U[P]. 2017-07-26.
- [27] Smith S. Method and apparatus for automatic linear shift parallax correction for multi-array image systems: US7738017 B2[P]. 2010-06-15.
- [28] Liluashvili V, Kalayci S, Fluder E, et al. iCAVE: an open source tool for visualizing biomolecular networks in 3D, stereoscopic 3D and immersive 3D [J]. *GigaScience*, 2017, 6(8): 1-13.
- [29] Yamaji M, Hayashi T. Depth perception in stereo graphics using off-axis projection [J]. *Journal of Informatics*, 2005, 23: 103-127.
- [30] VanArsdale D. Homogeneous transformation matrices for computer graphics [J]. *Computers & Graphics*, 1994, 18(2): 177-191.
- [31] de Sorbier F, Nozick V, Saito H. Multi-view rendering using GPU for 3-D displays [J]. *GSTF International Journal on Computing*, 2010, 1(1): 44-49.
- [32] Zhou W P, Tang H, Ji Z. GPU instancing method for crowd animation based on motion captured data [J]. *Journal of Computational Information Systems*, 2013, 9(11): 4459-4467.
- [33] do Nascimento B T, Franzin F P, Pozzer C T. GPU-based real-time procedural distribution of vegetation on large-scale virtual terrains [C] // 2018 17th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames), October 29-November 1, 2018, Foz do Iguacu, Brazil. New York: IEEE Press, 2019: 18434878.
- [34] Wade N J, Hughes P. Fooling the eyes: trompe l'oeil and reverse perspective [J]. *Perception*, 1999, 28(9): 1115-1119.
- [35] Cook N D, Hayashi T, Amemiya T, et al. Effects of visual-field inversions on the reverse-perspective illusion [J]. *Perception*, 2002, 31(9): 1147-1151.
- [36] Munkberg C J, Akenine-Möller T G, Salvi M, et al. Backface culling for motion blur and depth of field [J]. *Journal of Graphics, GPU, and Game Tools*, 2010, 15(2): 123-139.
- [37] Nyberg D. Analysis of two common hidden surface removal algorithms, Painter's algorithm & Z-Buffering [EB/OL]. [2021-03-15]. https://xueshu.baidu.com/usercenter/paper/show?paperid=20405c9bcfa94cce386cf9b278b0fdb6&site=xueshu_se.
- [38] Purcell T J, Buck I, Mark W R, et al. Ray tracing on programmable graphics hardware [J]. *ACM Transactions on Graphics*, 2005, 21(3): 703-712.
- [39] Xing S J, Sang X Z, Yu X B, et al. High-efficient computer-generated integral imaging based on the backward ray-tracing technique and optical reconstruction [J]. *Optics Express*, 2017, 25(1): 330-338.
- [40] Huo W P, Sang X Z, Xing S J, et al. Backward ray tracing based rectification for real-time integral imaging display system [J]. *Optics Communications*, 2020, 458: 124752.
- [41] Parker S G, Robison A, Stich M, et al. OptiX: a general purpose ray tracing engine [C] // ACM SIGGRAPH 2010 papers on-SIGGRAPH '10, July 26-30, 2010, Los Angeles, California. New York: ACM Press, 2010: 1-10.
- [42] Miyazawa T, Koyamada K. A high-speed integrated renderer for interpreting multiple 3D volume data [J]. *The Journal of Visualization and Computer Animation*, 1992, 3(2): 65-83.
- [43] Ruijters D. Dynamic resolution in GPU-accelerated volume rendering to autostereoscopic multiview lenticular displays [J]. *EURASIP Journal on Advances in Signal Processing*, 2008, 2009: 843753.
- [44] Pang B, Sang X Z, Xing S J, et al. High-efficient rendering of the multi-view image for the three-dimensional display based on the backward ray-tracing technique [J]. *Optics Communications*, 2017, 405: 306-311.
- [45] Li Y H, Sang X Z, Xing S J, et al. Real-time optical 3D reconstruction based on Monte Carlo integration and recurrent CNNs denoising with the 3D light field display [J]. *Optics Express*, 2019, 27(16): 22198-22208.
- [46] Huo W, Cheung G, Xin L, et al. Depth map super-resolution using synthesized view matching for depth-image-based rendering [C] // IEEE International Conference on Multimedia & Expo Workshops, July 9-13, 2012, Melbourne, VIC, Australia. New York: IEEE Press, 2012: 12935944.
- [47] Bravo G, Do L, Zinger S, et al. Real-time free-viewpoint DIBR on GPUs for 3DTV systems [C] // 2011 IEEE International Conference on Consumer Electronics-Berlin (ICCE-Berlin), September 6-8, 2011, Berlin, Germany. New York: IEEE Press, 2011: 1-4.
- [48] Wang L H, Hou C P, Lei J J, et al. View generation with DIBR for 3D display system [J]. *Multimedia Tools and Applications*, 2015, 74(21): 9529-9545.
- [49] Fan Y C, Chen Y C, Chou S Y. Vivid-DIBR based 2D-3D image conversion system for 3D display [J]. *Journal of Display Technology*, 2014, 10(10): 887-898.
- [50] Li S, Zhu C, Sun M T. Hole filling with multiple reference views in DIBR view synthesis [J]. *IEEE Transactions on Multimedia*, 2018, 20(8): 1948-1959.

- [51] Yao L, Han Y D, Li X M. Virtual viewpoint synthesis using CUDA acceleration[C]//Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology, November 2, 2016, Munich, Germany. New York: ACM, 2016: 367-368.
- [52] Lin J P, Wang W X, Yao J M, et al. Fast multi-view image rendering method based on reverse search for matching[J]. Optik, 2019, 180: 953-961.
- [53] Xing S J, Sang X Z, Cao L C, et al. A real-time super multiview rendering pipeline for wide viewing-angle and high-resolution 3D displays based on a hybrid rendering technique[J]. IEEE Access, 2020, 8: 85750-85759.
- [54] Muftuler L T, Nalcioglu O. Improvement of temporal resolution in fMRI using slice phase encode reordered 3D EPI [J]. Magnetic Resonance in Medicine, 2000, 44(3): 485-490.
- [55] Halle M. Multiple viewpoint rendering [C] // SIGGRAPH'98: Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM, 1998: 243-254.
- [56] Guan Y X, Sang X Z, Xing S J, et al. Parallel multi-view polygon rasterization for 3D light field display [J]. Optics Express, 2020, 28(23): 34406-34421.
- [57] van Berkel C. Image preparation for 3D LCD [J]. Proceedings of SPIE, 1999, 3639: 84-91.
- [58] Lee S, Park J, Heo J, et al. Autostereoscopic 3D display using directional subpixel rendering [J]. Optics Express, 2018, 26(16): 20233.
- [59] Li D X, Zang D N, Qiao X T, et al. 3D synthesis and crosstalk reduction for lenticular autostereoscopic displays[J]. Journal of Display Technology, 2015, 11(11): 939-946.
- [60] Yamamoto K, Ichihashi Y, Senoh T, et al. Simplified homography matrix for a large-scale camera array system[J]. Applied Optics, 2011, 50(34): H203-H210.
- [61] Xing S J, Liu S, Sang X Z. Multi-projector three-dimensional display for 3D Geographic Information System[J]. Optik, 2017, 139: 385-396.
- [62] Ruppert R, Atashband F. UDP broadcast for user interface in a download and configuration gaming method: US8260946B1[P]. 2010-09-04.
- [63] Martin T. The NFS lite CE installation on OSG[EB/OL]. (2016-12-19)[2020-03-14]. https://www.cac.cornell.edu/wiki/index.php?title=OSG_Cluster.

Overview of Virtual Stereo Content Generation Technology for Super Multi-View Light Field

Xing Shujun^{1, 2}, Cao Liangcai², Sang Xinzhu¹, Yu Xunbo¹, Jin Guofan²

¹ State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications (BUPT), Beijing 100876, China;

² State Key Laboratory of Precision Measurement Technology and Instruments, Tsinghua University, Beijing 100084, China

Abstract

Significance Recently, super multi-view light-field display devices have made great progress, showing the characteristics of increasing views, higher resolution, and larger viewing angles. The virtual scene to be displayed is becoming increasingly complex.

Herein, we review and summarize the existing virtual stereo content generation technology for super multiview light-field display devices. Their applicable scenarios, merits, and defects are illustrated. The main challenges of content generation of existing super multiview light-field display devices are highlighted. Real-time light-field content generation technology for large-scale virtual scenes of super multiview is not fully mature. A new rendering system needs to be explored to break through the bottleneck of existing algorithms. With the technical breakthrough, the popularization and application of light-field in key fields such as military affairs, agriculture, and fire protection could be solved.

Progress Recently, the super multiview virtual content generation technology has made consistent progress in the following three aspects, which are summarized in table 2.

Graphic processing unit (GPU) instancing is used to improve the view-by-view light-field rendering technology (Fig. 18). When there are many repetitive objects in the virtual scene that needs to be drawn, the GPU instancing

technology only requires to set the original data buffer once and then draw calls. Compared with the original view-by-view rendering algorithm based on geometric shaders, as the number of views increases, GPU instancing requires no more memory. When there is no object repeatedly drawn in the virtual scene, the super multiview light-field image of a large-scale scene can be rendered simultaneously.

The delay shading technique is introduced to improve the super multiview light-field rendering technique based on depth rendering (Fig. 23). Depth-based calculations can produce false illumination information, which is considerably different from the illumination characteristic in real space. The three-dimensional technology team of the Beijing University of Posts and Telecommunications proposed an image generation method of super multiview light-field based on delay coloring in 2020. Its reference image comprises one or more pairs of color, depth, highlight, and environment images. Although the lighting problem is solved using this method, it still leads to the degradation of multiview image quality and depth-based rendering.

An ultrafast experimental light-field rendering pipeline is constructed using geometric correlation (Fig. 24). In 2020, a super multiview rendering pipeline for polygon rasterization based on the principle of geometric correlation was developed. In the case of a 2×10^8 surface slice with a total resolution of 7680×4320 , the rendering rate reached 60 frame/s. Other factors have not been considered, such as lighting model and texture mapping, so it cannot be compatible with the traditional hardware rendering pipeline.

Conclusions and Prospect Many rendering algorithms can be used to generate super multiview light-field content. The ray tracing algorithm has the highest rendering quality, which can be used for nonreal-time applications. The algorithm with the fastest speed is the rendering algorithm based on geometric correlation; however, it is incompatible with the existing hardware rendering pipeline. Therefore, that lighting and occlusion need to be treated specially in the rendering process. Currently, the two most practical real-time rendering technologies are the rendering algorithm based on GPU instancing and the deferred rendering method. The former could degrade the image quality, whereas the latter is not fully competent for high-quality rendering of large-scale scenes. In general, the real-time light-field content generation technology for large-scale virtual scenes of super multiview is not yet fully developed. Therefore, it is still necessary to explore a new rendering system to break through the bottlenecks of existing algorithms.

Key words image processing; light-field display; ultra-multi-view rendering; virtual content generation; real-time rendering

OCIS codes 100.6890; 120.2040