

文章编号: 2095-2163(2023)11-0022-08

中图分类号: TP751

文献标志码: A

BVSL-34: 面向光场图像处理任务的基线可变仿真光场数据集

张宇翔¹, 吴天昊¹, 鲜明¹, 张洪江²

(1 国防科技大学 电子科学学院, 长沙 410073; 2 安康学院 电子与信息工程学院, 陕西 安康 725000)

摘要: 伴随着计算机视觉算法的迅猛发展, 为基于光场数据集的图像处理任务带来了广泛的关注, 数据是算法的基础, 光场数据集的制作正逐渐成为研究的重点问题。目前光场成像系统有多种实现途径, 其中光场阵列相机作为新体制的光场成像系统广受关注与研究, 但是目前基于此法制作的主流光场数据集并不能提供场景基线可变化的需求, 极大地限制了处理任务的预期效果。为此, 本文面向相关光场图像处理任务研究需求, 基于 3dsMax 软件构建了一种全新的基线可变化的仿真光场数据集 (Baseline-variable simulated lightfield-34, BVSL-34), 解决了主流数据集中不能针对阵列相机基线变化的缺陷, 该数据集包含 3 个大类, 共计 34 个场景, 534 个实例, 且各场景配备基线可变化的阵列相机。本文详细阐述了数据集的制作方法, 并与多个主流数据集进行性能、指标对比, 同时进行光场图像重聚焦实验, 结果验证了 BVSL-34 数据集功能全面、质量优异, 可以将该方法应用于相关光场图像处理任务中。

关键词: 光场数据集; 三维重建; 3dsMax; 图像处理

BVSL-34: A baseline variable light field dataset for light field image processing tasks

ZHANG Yuxiang¹, WU Tianhao¹, XIAN Ming¹, ZHANG Hongjiang²

(1 College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China;

2 School of Electronics and Information Engineering, Ankang University, Ankang Shaanxi 725000, China)

Abstract: With the rapid development of computer vision algorithms, it has brought widespread attention to image processing tasks based on light field datasets. Data is the basis of algorithms, and the production of light field datasets is gradually becoming a key issue of research. At present, there are many ways to realize the light field imaging system. Among them, the light field array camera is widely concerned and studied as a new light field imaging system. However, the current mainstream light field datasets based on this method cannot provide the scene baseline variable, greatly limiting the expected effect of processing tasks. To this end, this paper builds a new baseline-variable simulated lightfield-34 (BVSL-34) based on the 3dsMax software for the research needs of related light field image processing tasks, which solves the problem of mainstream data. Focusing on the defects that cannot be focused on the variation of the array camera baseline, the dataset contains 3 categories, a total of 34 scenes, and 534 instances, and each scene is equipped with an array camera with a changeable baseline. This paper expounds the production method of the data set in detail, and compares the performance and indicators with several mainstream data sets. At the same time, the light field image refocusing experiment is carried out. The results verify that the BVSL-34 data set has comprehensive functions and excellent quality, which could be applied to related light field image processing tasks.

Key words: light field dataset; 3D reconstruction; 3dsMax; image processing

0 引言

图像记录了光的信息, 也是人类记录三维世界的手段。纵览人类历史, 实际上也是人类对于光学

信息的研究发展史, 从简笔绘画到精美的写实画, 从手工绘制地图到如今的数字实景地图。然而目前的相机成像技术仍依赖于针孔成像模型, 是使用传感器记录空间中某一平面的光强二维分布, 是对七维

基金项目: 国家自然科学基金(61972435, 61401474)。

作者简介: 张宇翔(1998-), 男, 硕士研究生, 主要研究方向: 深度学习、信息安全; 吴天昊(1998-), 男, 硕士研究生, 主要研究方向: 计算成像、深度学习; 张洪江(1975-), 男, 讲师, 主要研究方向: 网络信息技术、数字媒体。

通讯作者: 鲜明(1970-), 男, 博士, 研究员, 博士生导师, 主要研究方向: 深度学习、信息安全。Email: 1069930599@qq.com

收稿日期: 2022-11-08

全光函数进行二维采样的过程。在传统成像技术的基础上, 计算摄影学从记录多维光学信息的角度, 克服了传统成像方法的以下局限:

(1) 景深和光圈大小的局限性。传统光学成像的镜头光圈限制成像的景深^[1]。光圈越大, 镜头对光信号的采集能力越强, 成像质量越好。但与之同时出现的问题是景深范围小, 不利于全场景成像。

(2) 二维采样的局限性。针孔成像模型中的图片为三维世界坐标系的点在二维上的投影, 因此传统成像模型无法采集深度信息。

(3) 镜头成本的局限性。传统相机要想获得更好的成像质量, 需要更大体积、更大光圈的镜头, 费用较为高昂^[1]。

在计算摄影学中, 近十多年逐渐成熟的光场成像(Light Field Imaging)已经成为现代光学成像的新兴领域。光场成像设备可以同时记录光线的强度和方向, 人们利用这一点可以提取出场景的三维信息。利用三维信息, 光场数据可以生成场景深度图。深度图像是三维场景重建的基础, 利用此信息可以实现更多复杂算法。此外, 光场成像还可以实现光场重聚焦^[2-3]、光场超分辨等功能。

在计算机视觉中, 可靠的数据是算法训练的基础。目前, 知名的公开光场数据集有斯坦福大学利用真实阵列相机建立的光场数据集 STF-gantry, Lytro 官网公开的数据集 EPFL 以及 INRIA。这些数据集都具有场景内容丰富、成像质量高的特点。真实数据集的优势在于拍摄的场景全部源于真实世界, 但是真实数据集依赖于现实条件, 例如昂贵的相机设备、还有拍摄的环境需要仔细选择, 否则成像质

量难以保证。而 HCI 数据集提供了一种新思路, 利用现在成熟的 3D 建模技术, 在电脑中先构建虚拟场景, 再通过渲染方式生成光场图片。由于在虚拟场景中生成图片是严格按照针孔成像模型进行 3 维到 2 维的映射, 这种方式得到的光场图片免去了繁琐的校正步骤, 而且由于整个场景的人工搭建, 还可以渲染出深度图作为深度的真值, 这是真实场景数据集做不到的。

然而, 目前主流的仿真数据集还有一个未能实现的功能是调节阵列相机基线, 因此本文基于 3dsMax 建模软件构建了一个基线可变的原创光场数据集—BVSL-34, 还开创性加入了对于同一场景的可变基线的功能, 并且后文对于光场算法的实现也将基于此数据集进行。

1 研究背景

1.1 光场阵列相机

阵列相机的原理是在四维光场中, 在多个位置放置相机对四维光场进行二维切片, 相当于对光场进行多角度的二维采样。

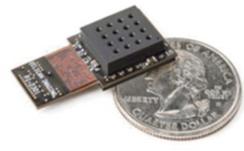
文献[4]设计了一个 8×8 的阵列相机。这种设备的特点是能够对运动的光场进行感知。文献[5]设计了一个 6×8 的光场成像设备, 可以渲染不在 6×8 视角位置的视角图像。就是通过对各子相机位置视角进行重排列, 通过这种手段获得更优的成像品质。文献[6]制作了一个 8×12 的阵列相机, 并且将获取的光场数据制作成公开数据集。各种体制阵列相机如图 1 所示。



(a) 西北工业大学阵列相机



(b) 斯坦福大学阵列相机



(c) PiCam 阵列相机

图 1 各种体制阵列相机

Fig. 1 Various institutional array cameras

来自西北工业大学的研究团队, 研制了一种 8×8 的相机阵列^[7], 见图 1(a)。他们基于这种阵列相机在相机校正^[8]等课题进行了深入研究。以上提到的阵列相机均为大型的光场成像设备, 价格昂贵。另一方面, 随着相机逐渐向小型化发展, 也有团队将

阵列相机的小型化作为目标。PiCam 是一款超薄 4×4 的阵列成像设备, 整体体积与硬币相当, 由 Venkataraman 的团队研制^[9], 见图 1(c)。综上, 阵列相机具有如下特点:

(1) 阵列相机可以瞬时获取光场。

(2)阵列相机可以大型化、也可以小型化,并且在成像设备集成度越来越高的情况下,小型化的阵列相机也有不俗表现。

1.2 3ds Max

3D Studio Max,常简称为 3d Max 或 3ds MAX,是 Discreet 公司开发的(后被 Autodesk 公司合并)基于 PC 系统的 3D 建模渲染和制作软件。其前身是基于 DOS 操作系统的 3D Studio 系列软件,在设计 3D 模型的 2D 横截面形状中发挥着巨大作用。

3ds Max 软件由一些重要的要素组成,包括主工具栏、命令面板、Maxscript 编辑器、视口、材质编辑器和渲染框窗口。主工具栏位于窗口顶部,包括选项和工具;命令面板位于窗口界面的右侧,包括了操作者需要的所有命令;Maxscript 编辑器作用是对脚本、实用程序和插件进行编码操作;视口是从上、左、前、透视图这 4 个不同角度显示场景的视图框,通过视口使得操作 3ds Max 工作更加轻松快捷;材质编辑器的作用是对场景对象单元进行纹理处理,并实现有关对象单元的材质和贴图操作;渲染框窗口是查看最终渲染场景的视图窗,支持用户使用材质和照明查看其场景。由于 3ds Max 具有一套便捷实用的命令、工具以及内置基础形状,例如圆锥形、椭球形、立方体、茶壶形、金字塔形等,极大地提升了建模效率和操作体验。同时,3ds Max 加入了对头发、皮肤、毛皮和衣服的仿真,这些内置工具减少了角色建模所需的时间并增强了细节。另外,3ds Max 还支持 NURBS 功能,使操作者可利用数学公式进行平滑建模。

2 数据集制作原理

3ds Max 是一种三维建模软件,具有建模速度快、功能丰富的特点,并且结合主流渲染器 V-ray,可以做到照片级渲染,在城市渲染图和建筑仿真制作中应用广泛。而 3ds Max 可以做到场景搭建、模拟真实光照以及场景内多视角观测,使得 3ds Max 更加适合用来制作阵列相机的光场仿真图像。本文使用该软件进行仿真数据集制作,制作流程如图 2 所示。



图 2 数据集制作流程图

Fig. 2 Flow data set production

本文提出的制作方法共分为 5 个步骤:

(1)在 3ds Max 软件中进行特定三维场景建模,定义出所有物体单元的形状与尺寸。

(2)对场景中的物体单元进行配置,赋予各个场景中单元材质与物理属性。

(3)在已构建场景中添加光照并进行参数调节。

(4)在已构建场景中的不同位置架设虚拟相机,并且搭建不同基线的虚拟相机阵列。

(5)利用已搭建完毕的虚拟阵列相机进行渲染成像,以实现光场仿真图像的制作。

以下将对各个步骤进行阐述。

2.1 三维场景建模

首先需要在 3ds Max 中搭建三维场景,此过程要精确搭建每一个三维物体的空间配合关系,并且指定物体的长宽高等基本空间物理属性,一次搭建即可以实现三维场景的多次快速复用。图 3 展示了在 3ds Max 软件中搭建场景的界面。

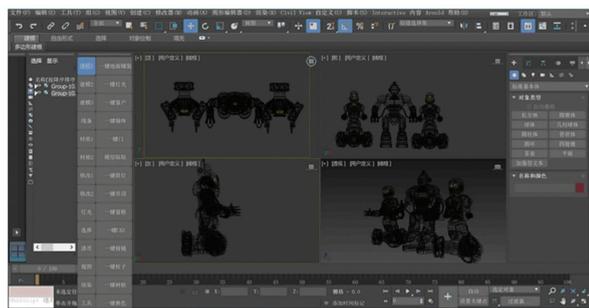


图 3 三维场景建模

Fig. 3 3D scene modeling

2.2 场景单元配置

场景构建完毕后,需要对三维场景中的物体单元进行贴图操作,同时为每个物体单元赋予相关材质和物理属性,包括物体反射系数、折射系数、颜色等固有属性。图 4 展示了场景单元参数配置的相关操作。

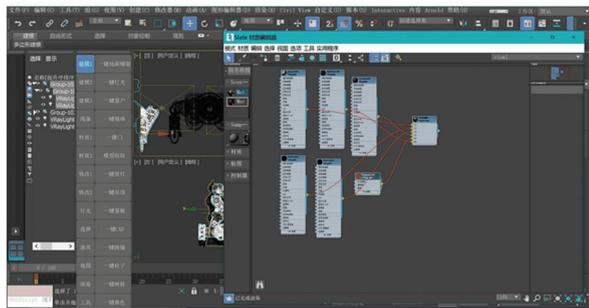


图 4 场景单元参数配置

Fig. 4 Scene unit parameter configuration

2.3 光照参数调节

为了更贴近现实世界,在模拟三维环境中添加光照调节是关键的一步,图 5 展示了在已构建的三维场景中进行光照参数调节的相关操作。通过设置光源的位置、形状、指向、光照强度等一系列参数以添加光源,同时可以便捷、快速地进行参数调整以满足符合特定环境要求的光照条件。

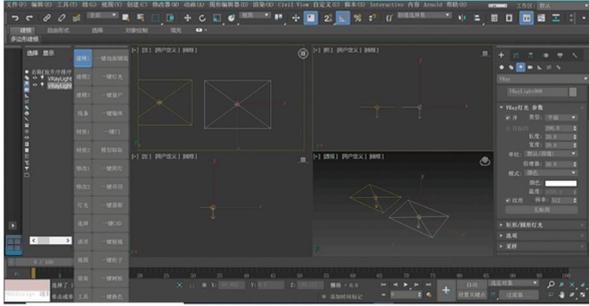


图 5 场景光照参数调节

Fig. 5 Scene lighting parameter adjustment

2.4 虚拟阵列相机搭建

场景构建及相关配置完毕后,需要在虚拟场景中的不同位置放置用于成像的虚拟阵列相机。根据场景光场图像成像的要求,在指定位置搭建不同基线的虚拟阵列相机。图 6 展示了在 3ds Max 中搭建虚拟阵列相机的界面,图中白色阵列结构即为虚拟阵列相机。

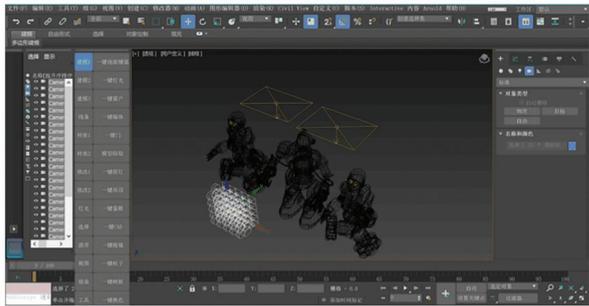


图 6 建立虚拟阵列相机

Fig. 6 Build a virtual array of cameras

2.5 渲染成像

场景中的虚拟阵列相机搭建完毕后,就可开始进行光场图像的渲染成像,这是整个制作过程中的核心步骤。在不同基线情况下,对场景进行渲染,本质是对于每一个阵列相机的子相机获得的场景进行渲染,渲染的过程实际上是仿真相机成像的过程。最后得到所需的光场仿真图像。图 7 展示了基线为 5 mm 的一组场景,图片的排布与实际阵列相机的分布一致。

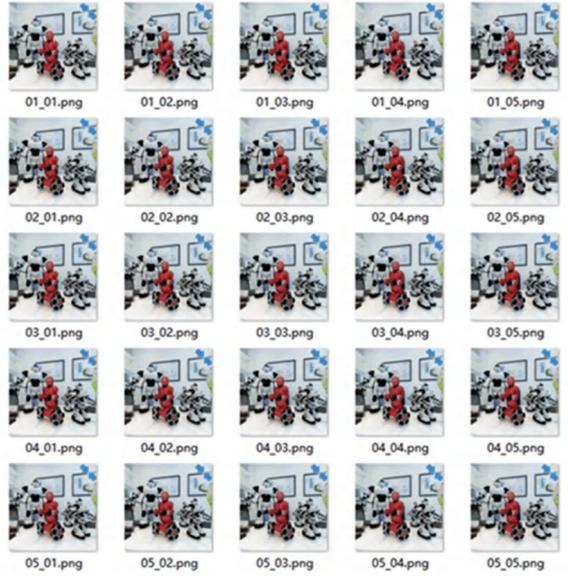


图 7 基线为 5 mm 的一组光场仿真图像

Fig. 7 A set of light field simulation images with a baseline of 5 mm

3 数据集展示

3.1 数据集特点

BVSL-34: 基线可变化的仿真光场数据集 (Baseline-variable simulated lightfield-34) 有以下特点:

(1) 阵列相机参数一致性完全可控。为了阵列相机成像质量的稳定性,往往需要阵列中的每个子相机具备完全相同的参数,显然在现实世界中,这种绝对意义上的精准与相同是无法达到的。但是利用本文提出的方法,在构建场景中搭建的虚拟阵列相机可以实现严格意义上的精准,子相机的等距阵列分布以及相机参数都可以做到完全相同。

(2) 理想化渲染成像。在 3ds Max 中构建的场景中,通过搭建的虚拟阵列相机可实现理想化的渲染成像。由于虚拟阵列相机成像模型完全按照理想针孔成像模型进行渲染成像,相比于现实世界中的阵列相机,省去了相机校正和去畸变的过程,提高了光场图像的成像质量,在后续进行算法验证时减少了预处理操作,提供了更加快捷的方式,提升了整体效率。

(3) 相机基线可调节。相较于传统光场图像制作方式,该方法最突出的优势是开创性加入了对于相同场景下可变基线的成像功能,可以实现场景中同一角度的阵列相机基线可变地成像,满足了场景基线可变化的需求。图 8 很好地展示了该方法对于仿真场景如何实现基线变化,图 8 中使用 3×3 的阵

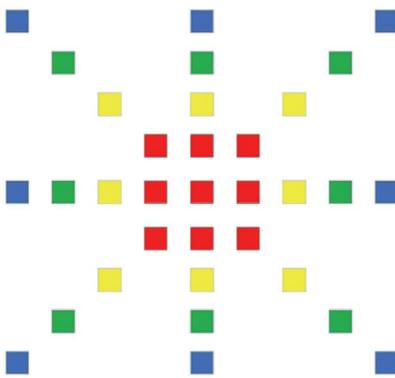
列相机作为展示,其中不同基线的阵列相机使用不同的颜色进行标注。

(4)具备深度图渲染能力。该方法同时具备深度图渲染能力,可同时生成场景图与对应的深度图,如图9所示,其中深度图可作为深度的真值图像使用。与传统光场深度图制作方式需要进行深度估计

不同,该方法制作深度图的原理是直接基于对三维场景的深度采样得到。这是因为在进行三维建模时,所有物体的三维信息都是精确确定的,并且在场景中均可直接观测到。图9(a)和图9(b)对比表明,生成的深度图与原场景完全贴合,并且具备显示深度全部细节的能力。



(a) 场景中虚拟阵列相机



(b) 不同基线示意图

图8 阵列相机场景基线可变示意图

Fig. 8 Schematic diagram of the variable baseline of the array camera scene



(a) 场景图



(b) 深度图

图9 场景图与对应深度图

Fig. 9 Scene graph and corresponding depth map

3.2 性能对比

本数据集具有34个场景、覆盖3大类,每个场景均进行了多基线的渲染,具有多种基线可供选择。3个大类分别为:自然风景、微距物体、室内场景,大类的选择是依据尽可能覆盖生活中拍摄可能出现的典型深度。自然风景为大场景、最远深度可达百米;微距物体则是小场景范围、最远深度在1~2 m;室内场景则为二者折中。图10依据场景展示了该数据集的部分数据。

下面是针对本数据集和其他主流光场数据集的质量对比,对比范围选取BRISQUE、NIQE、CEIQ和ENIQA指标。本数据集和其他主流光场数据集特点对比结果见表1。本数据集和其他主流光场数据

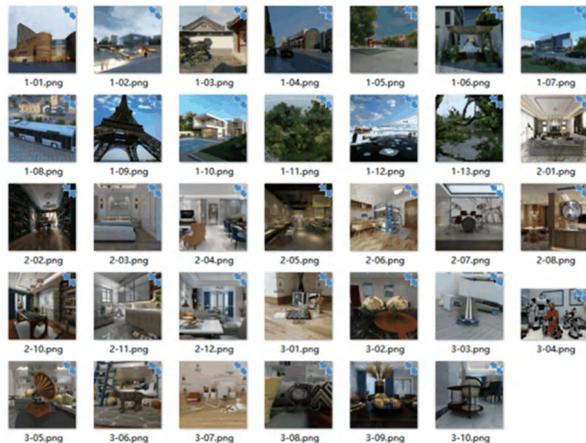


图10 部分数据集展示

Fig. 10 Part of the dataset

集关键指标对比结果见表 2。由表 1 和表 2 可知, 本文的数据集具有良好的性能表现, 表 2 中的加粗数据为最优结果。

表 1 本数据集和其他主流光场数据集特点对比

Tab. 1 Comparison of the characteristics of this dataset and other mainstream light field datasets

数据集	数据集类型	场景数	深度图	基线
HCInew ^[19]	仿真	24	有	固定
HCIold ^[20]	仿真	12	有	固定
STFgantry ^[21]	真实	12	无	固定
BVSL	仿真	34	有	可变

表 2 本数据集和其他主流光场数据集关键指标对比

Tab. 2 Comparison of key indicators between this dataset and other mainstream light field datasets

数据集	BRISQUE(↓)	NIQE(↓)	CEIQ(↑)	ENIQA(↓)
HCInew ^[19]	14.800	3.833	3.153	0.087
HCIold ^[20]	24.170	2.985	3.369	0.117
STFgantry ^[21]	25.280	4.246	2.781	0.232
BVSL	8.901	3.593	3.375	0.041

可见, 本数据集在 4 个数据指标中有 3 个领先、1 个次优, 在各种功能综合表现上也处于全面领先, 场景数量多, 具有深度数据, 能进行基线变化。

4 数据集仿真图像验证

重聚焦是光场相机对比传统成像设备的特有功能, 重聚焦技术是指光场成像设备获取光场数据后, 利用重聚焦算法, 将成像的图片调焦到场景中的任意位置, 利用重聚焦技术可以实现“先拍照后聚焦”的功能, 颠覆了传统的拍照成像模式。因此, 本节将利用光场图像具备重聚焦这一特性, 采用基于重投影的阵列相机重聚焦实验来对数据集仿真图像进行测试, 验证本数据集的光场图像是否符合光场图像处理任务的需求。

4.1 实验方法

(1) 基于重投影的阵列相机重聚焦方法。2000 年, 文献[10]提出的动态参数法可以将光场数据聚焦到不同的深度。2004 年, 文献[11]基于单应性方法提出: 在相机校正得到外参数的前提下, 将光场图像反投影到人为指定的参考平面上, 对参考平面上的子图像重投影到参考相机平面进行平移重叠相加, 获取聚焦在指定深度参考平面上的重聚焦图像。

该方法流程及原理如图 11 所示。利用重投影思想将光场图像反投影到一个指定深度的参考平面

上, 对参考平面上的子图像重投影到参考相机平面进行平移重叠相加, 获取聚焦在指定深度的重聚焦图像。在聚焦过程中, 对于没有严格对齐的像素位置采取一次插值方法获取插值图像。相较于多次插值方法, 一次插值可以尽可能减少插值产生的误差。



图 11 重聚焦算法流程示意图

Fig. 11 Flow diagram of refocusing algorithm

(2) 重聚焦。重聚焦是光场成像方式中一种基本方法。核心是使用“三维空间中不同位置的物体在不同位置的子相机成像中产生的视差 (disparity) 信息偏差”理论。将光场图像投影到一个参考平面上, 对参考平面上的子图像进行不同位置的位移后求和, 取均值获取聚焦在指定深度的重聚焦图片。图 12 展示了双目相机对同一物体的不同视角的视差区别及聚焦和虚化原理。

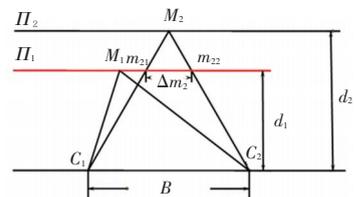


图 12 双目重聚焦原理示意图

Fig. 12 The schematic of binocular refocusing principle

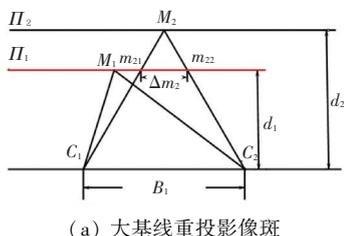
图12中,双目相机从2个视角获取的图像信息具有视差,处于聚焦平面的物体在2个视角的图片中视差为0,即是聚焦。而不在聚焦平面上的点,由于具有视差,在图像叠加过程中就会产生模糊,也就是摄影学中常说的虚化效果。

(3)重投影。重投影是将每个相机投影在聚焦面上的三维点重新投影到其他相机(通常是中心参考相机)的像平面上。设三维点坐标仍为 (x_w, y_w, z_w) ,那么该点代入第 r 个相机的成像模型中,在第 r 个相机像平面上的像素坐标为 (u_r, v_r) ,可得:

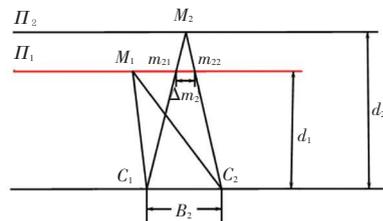
$$s \begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} = K_r [R_r \quad T_r] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (1)$$

得到了每个相机重新投影到中心参考相机像平面的重投影图像,聚焦就是将这些重投影图像进行叠加取均值。为此可推得:

$$I = \frac{1}{N} \sum_{i=1}^N I'_i \quad (2)$$



(a) 大基线重投影像斑



(b) 小基线重投影像斑

图13 不同基线对重投影像斑的影响

Fig. 13 Effect of different baselines on reprojected image spots



(a) 未重聚焦中心视角

(b) 2.5 mm 基线重聚焦

(c) 5 mm 基线重聚焦

图14 相机基线对于光场重聚焦结果的影响

Fig. 14 Effect of camera baseline on light field refocusing results

实验结果表明,2组光场仿真图像重聚焦后均聚焦在床架上,在相机基线为2.5 mm时,重聚焦的景深范围较大,前后景略有虚化;在相机基线为5 mm时,景深范围比2.5 mm场景小,前后景虚化比2.5 mm更明显,实验结果与理论预期一致。

对于在聚焦平面上的点,所有相机的反投影点都在同一点,经过重投影仍然汇聚在同一个像素点;对于不在聚焦平面的点,所有相机的反投影点并不重合,因此重投影会形成一个模糊的范围,也就是焦外虚化。经过求均值,最终便得到光场图像的重聚焦图像。

4.2 验证与分析

使用基于重投影的阵列相机重聚焦方法进行验证,实验数据采用按照本文方式制作的光场仿真图像,进行不同相机基线场景下的重聚焦算法测试与结果分析。

图13展示了理论上不同基线对重投影像斑的影响。图13中, Δm_2 为非聚焦平面上的点在重聚焦时产生的重投影像斑的直径,对于同一个点 M_2 ,大基线阵列相机产生的重投影像斑大于小基线阵列相机产生的重投影像斑。

实验选择2组不同基线的光场仿真图像进行重聚焦实验,实验结果如图14所示,展示了同一场景下不同基线的重聚焦结果。

整个实验过程及结论表明,本文提出的基于3ds Max三维重建的光场图像仿真制作方法所生成的光场仿真图像具备光场图像的特有属性,完全符合光场图像实验要求。同时,与传统方法相比,实际的相机获取的光场数据在进行光场算法实验时需要

考虑畸变, 而通过本文方法制作的光场数据在实验过程中完全不必考虑这一复杂的畸变处理过程。另外, 本方法还实现了同一场景中阵列相机基线可变动地成像, 满足了同一场景基线可变化的实验需求, 且图像质量均衡、自动化程度高, 能够有效提高特定场景下光场图像的制作效率。

5 结束语

本文构建了一个基线可变的仿真光场数据集 BVSL-34, 包含 3 个大类, 共计 34 个场景, 534 个实例, 并详细介绍了数据集的制作方法, 该方法解决了传统光场图像制作中存在的过于依赖现实世界条件、成像质量难以保证、阵列相机基线不可变等难点问题, 极大地拓宽了光场图像的应用场景。与现有主流数据集 HCInew^[12]、HCIdold^[13]、STFgantry^[14] 相比, BVSL-34 数据集具有相机参数稳定, 成像基线可变, 同时呈现深度图等优势特点, 对比显示, 本数据集在成像质量以及各项相关指标上均表现优异。最后, 基于重投影的阵列相机重聚焦实验验证了 BVSL-34 数据集仿真图像具备光场数据的特性, 整体制作与实验流程效果较传统方法有显著的优势, 结果验证了 BVSL-34 数据集功能全面且质量优异, 可以将该方法应用于相关光场图像处理任务中。

参考文献

- [1] SCHECHNER Y Y, KIRYATI N. Depth from Defocus vs. Stereo: How Different Really Are They? [J]. *International Journal of Computer Vision*, 2000, 39(2): 141-162.
- [2] VAISH V, WILBURN B, JOSHI N, et al. Using plane+ parallax for calibrating dense camera arrays[C]// *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington D.C. USA: IEEE, 2004: 2-9.
- [3] ISAKSEN A. Dynamically reparameterized light fields [C]// *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*. New York: ACM, 2000: 297-306.
- [4] YANG J C, EVERETT M, MCMILLAN B C. A real-time distributed light field camera [C]// *Eurographics Workshop on Rendering*. [S.l.]: Eurographics Association, 2002: 77-86.
- [5] ZHANG Cha, CHEN T. A self-reconfigurable camera array [C]// *Proceedings of the Fifteenth Eurographics Conference on Rendering Techniques*. New York: Eurographics Association, 2004: 243-254.
- [6] WILBURN B, JOSHI N, VAISH V, et al. High performance imaging using large camera arrays [J]. *ACM Transactions on Graphics (TOG)*, 2005, 24(3): 765-776.
- [7] XIAO Zhaolin, SI Lipeng, ZHOU Guoqing. Seeing beyond foreground occlusion: A joint framework for SAP-based scene depth and appearance reconstruction [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2017, 11(7): 979-991.
- [8] ZHANG Qi, ZHANG Chunping, LING Jinbo, et al. A generic multi-projection-center model and calibration method for light field cameras [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 41(11): 2539-2552.
- [9] VENKATARAMAN K, LELESCU D, DUPARRÉ J, et al. Picam: An ultra-thin high performance monolithic camera array [J]. *ACM Transactions on Graphics (TOG)*, 2013, 32(6): 166.
- [10] ISAKSEN A, MCMILLAN L, GORTLER S J. Dynamically reparameterized light fields [C]// *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*. New York: ACM, 2000: 297-306.
- [11] VAISH V, WILBURN B, JOSHI N, et al. Using plane+ parallax for calibrating dense camera arrays [C]// *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington DC, USA: IEEE, 2004: 1: 2-9.
- [12] HONAUER K, JOHANNSEN O, KONDERMANN D, et al. A dataset and evaluation methodology for depth estimation on 4d lightfields [C]// *Asian Conference on Computer Vision (ACCV)*. Taipei: Springer, 2016: 19-34.
- [13] WANNER S, MEISTER S, GOLDLUECKE B. Datasets and benchmarks for densely sampled 4d light fields [J]. *Vision, Modelling and Visualization (VMV)*, 2013, 13: 225-226.
- [14] VAISH V, ADAMS A. The (new) stanford light field archive [Z]. California: Stanford University, 2008.
- [15] Stanford (New) Light Field Archive. [EB/OL]. <http://lightfield.sstanford.edu/lfs.html>
- [16] UNGER J, WENGER A, HAWKINS T, et al. Capturing and rendering with incident light fields [C]// *EGRW '03: Proceedings of the 14th Eurographics workshop on Rendering*. Goslar, Germany: ACM, 2003: 141-149.
- [17] DANSEREAU D G, SCHUSTER G, FORD J, et al. A wide-field-of-view monocentric light field camera [C]// *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA: IEEE, 2017: 3757-3766.
- [18] POLLEFEYS M, Van GOOL L, OOSTERLINCK A. The modulus constraint: A new constraint for self-calibration [C]// *International Conference on Pattern Recognition*. San Francisco, CA, USA: IEEE, 1996, 1: 349-353.
- [19] WRIGHT J, MA Yi, MAIRAL J, et al. Sparse Representation for Computer Vision and Pattern Recognition [J]. *Proceedings of the IEEE*, 2010, 98(6): 1031-1044.
- [20] LYTRO. 3D light-field vision [EB/OL]. [2008]. <https://www.lytro.com/>.
- [21] TRIGGS B. Autocalibration and the absolute quadric [C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Puerto Rico, USA: IEEE, 1997: 610-614.
- [22] ZELLER C, FAUGERAS O, ANTIPOLIS I S, et al. Camera self-calibration from video sequences: The Kruppa equations revisited [R]. Finland Ireland HAL-INRIA, 1996.
- [23] DANSEREAU D G, PIZARRO O, WILLIAMS S B. Decoding, calibration and rectification for lenselet-based plenoptic cameras [C]// *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Portland, USA: IEEE, 2013: 1027-1034.