

光场相机三维重建研究进展与展望

刘玉轩, 张 力, 艾海滨, 许 彪, 孙钰珊, 樊仲黎

(中国测绘科学研究院摄影测量与遥感研究所, 北京 100036)

摘要: 光场相机利用二维影像同时记录空间中光线的位置和方向信息, 能够恢复相机内部的光场, 为三维重建提供了新的解决思路. 围绕光场相机的三维重建问题, 本文综述了光场数据获取手段, 梳理和讨论了针对光场相机的标定算法, 总结和分析了基于光场影像的三维深度信息恢复方法. 在此基础上, 介绍了当前主要的公开光场数据和算法. 最后, 展望了未来的研究方向, 以为后续研究者提供参考.

关键词: 光场相机; 三维重建; 光场采集; 相机标定; 深度估计; 公开光场资源

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112(2022)07-1774-19

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20211615

Progress and Prospect of 3D Reconstruction Based on Light Field Cameras

LIU Yu-xuan, ZHANG Li, AI Hai-bin, XU Biao, SUN Yu-shan, FAN Zhong-li

(Institute of Photogrammetry and Remote Sensing, Chinese Academy of Surveying and Mapping, Beijing 100036, China)

Abstract: Light field cameras use two-dimensional images to simultaneously record the position and direction information of light in space, which can restore the light field inside the camera and provide a new solution for three-dimensional reconstruction. Focusing on the 3D reconstruction of light field cameras, this paper summarizes the light field acquisition methods, sorts out and discusses the calibration algorithms for light field cameras; then summarizes and analyzes the 3D depth information recovery methods based on light field images. In addition, the primary open-source datasets and algorithms related to light fields are introduced. Finally, the future research directions are elaborated to provide references for subsequent researchers.

Key words: light field cameras; 3D reconstruction; light field acquisition; camera calibration; depth estimation; public light field resources

1 引言

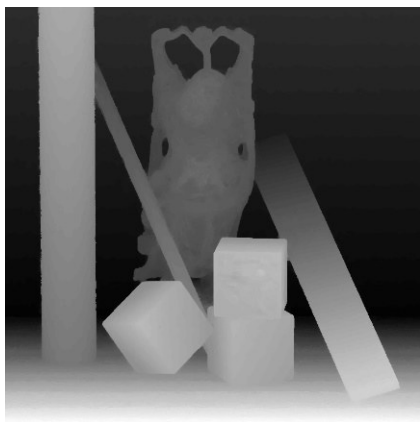
三维重建一直是摄影测量和计算机视觉的一个基本且重要的问题^[1,2]. 相比于二维图像, 三维模型包含了更加丰富的几何、形状和结构信息, 为场景理解与交互提供了更多可能. 具体地, 三维重建技术是让计算机感知周围的三维环境结构、物体的形状以及物体之间的相对位置等. 随着传感器技术的飞速发展, 三维重建数据呈现爆炸式增长, 三维重建技术也开始更多地走进人们日常生活的方方面面, 如自动驾驶^[3]、虚拟现实^[4]、增强现实^[5]、机器人^[6]、文物保护^[7]、人脸识别^[8]和医疗健康^[9]等.

目前, 研究人员在三维重建领域已经进行了诸多

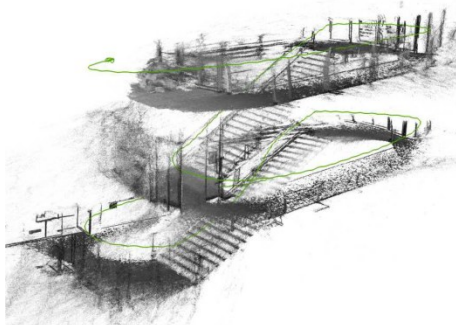
研究. 根据所用传感器的不同, 主要可分为基于传统影像的三维重建、基于激光点云的三维重建以及基于RGB-D相机的三维重建. 基于传统影像和激光点云的三维重建在理论和实践中都相对成熟, 比如基于传统影像的PTAM(Parallel Tracking And Mapping)^[10]、MonoSLAM^[11]、ORB-SLAM^[12]、ORB-SLAM2^[13]和ORB-SLAM3^[14], 以及基于激光点云的Cartographer^[15]、LOAM(Lidar Odometry And Mapping)^[16]、Lio-mapping^[17]、hdl_graph_slam^[18]、Fast-lio^[19]和SuMa++^[20]等. 基于RGB-D相机的三维重建虽然起步稍晚, 但也取得许多优秀的成果, 比如DVO^[21]、KinectFusion^[22]、ElasticFusion^[23]和BundleFusion^[24]、RobustFusion^[25]等. 尽管这些传感器在室内三维重建中得到越来越广泛的应用, 但

仍存在一些局限. 基于激光点云的三维重建虽然重建精度较高,但是重建出来的三维模型没有纹理信息. 单目传感器捕捉真实世界的二维影像,要实现三维重建必须对场景进行多次拍摄构建立体,但由于无法恢复场景的真实尺度,因此得到的只是相对尺度的三维模型. 大多数的 RGB-D 相机在自然光环境下表现较差,而且量程有限,随着量程的增加,精度会大幅度降低.

近年来,光场相机取得了较大的发展,已经可以得到与传统工业相机差不多的大小,甚至直接集成入手机^[26-30]. 与传统相机相比,光场相机不仅捕获空间点所有方向光线的累积强度,而且记录了到达某一空间点各个方向光线的强度大小,因此通过一次拍摄获取到多个视角的信息. 利用这些多视角信息,可以进行影像重聚焦、增大影像的景深以及评估场景的深度,且不易受外界环境的影响,为三维重建提供了新的可靠数据源. 图 1(a)给出了单张光场影像的深度估计结果^[31],图 1(b)则给出了利用多次拍摄光场影像恢复的包含绝对尺度的大场景半稠密三维模型^[32].



(a) 基于光场影像恢复的三维深度图^[31]



(b) 基于光场影像重建的半稠密三维点云模型^[32]

图 1 光场三维重建示例

高精度的光场相机三维重建不仅促进了传统三维技术的发展,而且推动了与光场相关领域技术的进步,比如光场超分重建^[33,34]、光场显著性检测^[35]、光场分

割^[36]、光场分类与识别^[37]和光场拼接^[38]等. 然而,由于光场相机的内部结构以及光场数据本身的特殊性,基于光场相机的三维重建依然存在较多问题^[39,40]. 首先,光场相机的内部结构完全不同于传统相机,导致其投影模型和相机畸变模型有别于传统相机. 其次,因为光场相机同时捕捉了空间中光线的方向和位置,所以一个空间点会在原始光场数据中成像多次,给光场相机的标定带来了难度. 最后,单次拍摄的光场数据,多视角的子孔径图像间的基线较短,传统的多视匹配算法应用于单次拍摄光场数据时效果较差.

针对以上问题,国内外在光场技术方面已经有了较多研究,本文对目前与光场三维重建相关的研究现状和算法,进行了梳理和总结,以期为后续研究人员提供参考和帮助.

2 光场的定义及采集

2.1 光场的概念及二平面表示

光场的概念最早是由 Gershun 于 1936 年提出^[41],它描述了空间中任意一点在任意方向光线的辐射亮度. 为了描述光场,Adelson 和 Bergaen 于 1991 年提出了全光函数^[42]. 全光函数包含了光线在空间中的所有基本特性,共 7 个维度:光线穿过某一点的坐标 (V_x, V_y, V_z) 、光线的方向 (θ, ϕ) 、光线的波长 λ 和到达某一点的时间 t ,可以被定义为

$$P(V_x, V_y, V_z, \theta, \phi, \lambda, t) \quad (1)$$

在大多数情况下,处理图像时并不关心获取影像时的绝对时间 t . 而且,单色传感器仅捕捉在一个较小的波段范围内的光线. 在这种情况下,光线独立于波长 λ ,因此式(1)中的七维函数可以减少为五维 $P(V_x, V_y, V_z, \theta, \phi)$. 对于 RGB 彩色传感器,可以把 3 个波段认为是 3 个分开的全光函数.

全光函数将光场描述为在空间和时间上连续的函数,但是捕获到的光场仅是整个光场中固定在光线传感器所在区域的部分. 虽然有多种方式可以用来表示采集到的光场数据,但其中最简单也最常用的方法是二平面法(Two Plane Parameterization, TPP)^[43]. 一方面是因为概念上的简单,另一个更重要的原因是它与目前大多数的光场采集设备具有很好的兼容性^[44]. 光场采集设备的主镜头和传感器是平行的平面,而且图像传感器上的像素是规则排列的,更容易利用 TPP 表示光场. 图 2 给出了 TPP 的示意图. 假设四维光场中一条光线同时穿过 2 个平行的平面 Ω 和 Π ,并且与两平面的交叉点分别为 $(u, v) \in \Omega$ 和 $(s, t) \in \Pi$,则光场 L 可表示为

$$L(u, v, s, t) \quad (2)$$

利用光线与2个平行平面的交点可以定义所有光线. 图2中 L_1 、 L_2 和 L_3 是从三维空间中一点 $O(V_x, V_y, V_z)$ 出的三条光线. 假如反射是朗伯体反射, 因为是从同一点射出, 则 L_1 、 L_2 和 L_3 的亮度值相等.

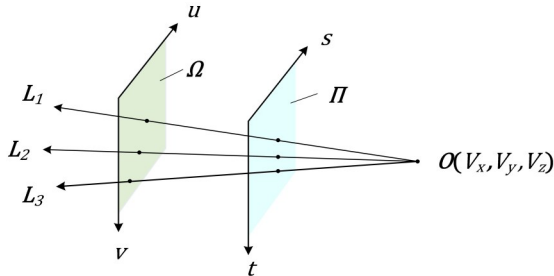


图2 光场的二平面表示

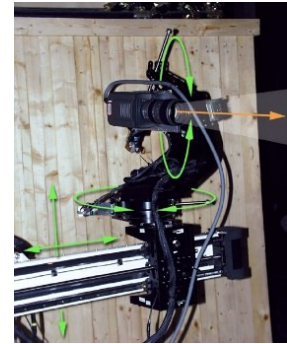
2.2 光场的采集

传统相机通过叠加各个方向到达同一像素的光线, 获得物方点在传感器平面上的二维投影. 光场获取设备或方法量测每条光线的位置和方向信息, 避免了不同方向光线的积分叠加, 从而捕获了每条光线的信息并恢复了设备内的光场. 根据构造的不同, 当前的光场获取设备主要分为4类: 多传感器捕获^[29, 45-49]、时序捕获^[50-52]、基于掩模^[53, 54]和基于微透镜阵列 (Micro-Lens Array, MLA)^[26, 38, 55-58]的光场相机. 图3给出了几种不同类型光场相机的示意图. 表1给出了当前的主要光场获取方法.

多传感器光场相机利用分布在平面或球面上的相机阵列同时捕获来自不同视角的光场信息, 相机的位置和拍摄的影像分别对应于TPP中的 (u, v) 和 (s, t) . Yang等^[45]于2002年利用 8×8 的视频相机捕捉了动态的光场. 同年, 斯坦福大学的Wilburn等^[46]利用6个CMOS



(a) 多传感器光场相机^[47]



(b) 基于时间序列拍摄的光场相机^[50]



(c) 基于掩膜的光场相机^[53]



(d) Lytro Illum^[56]

图3 四种典型的光场相机

图像传感器来同步记录视频数据获得光场. 随后, Wilburn等^[47]构建了如图3(a)所示的相机阵列来捕捉光场. Zhang和Chen^[48]利用 6×8 的相机阵列来记录光场, 以合成虚拟视角影像, 同时通过改变相机位置获得了更好的效果. 这些多传感器光场相机体积巨大, 为此, Venkataraman等^[29]设计了一个微型的光场相机PiCam. PiCam包含 4×4 的相机阵列, 每个相机可以捕捉 $1\,000 \times 750$ 像素的影像, 而且整个相机机身很小, 可以集成到手机. Lin等^[49]设计了包含 5×5 相机阵列的光场显微镜, 专门用于微小场景的拍摄.

与多传感器方法相比, 时序捕获方法使用单个图

表1 光场获取方法小结

光场相机类别	方法	时间	实现	分辨率	拍照速度
多传感器	文献[45]	2002	8×8 camera array	$320 \times 240 \times 8 \times 8$	15~20 fps
	文献[48]	2004	6×8 camera array	$320 \times 240 \times 6 \times 8$	15~20 fps
	文献[47]	2005	10×10 camera array	$640 \times 480 \times 10 \times 10$	30 fps
	文献[29]	2013	4×4 camera array	$1\,000 \times 750 \times 4 \times 4$	25 fps
	文献[49]	2015	5×5 camera array	$1\,024 \times 768 \times 5 \times 5$	30 fps
时序	文献[50]	2002	Gantry	$1\,300 \times 1\,030 \times 62 \times 56$	-
	文献[51]	2013	Linear stage	$5\,616 \times 3\,744 \times 100$	-
	文献[52]	2008	Programmable aperture	$3\,039 \times 2\,014 \times 5 \times 5$	0.5 s
掩膜	文献[53]	2007	Attenuation mask	$228 \times 181 \times 9 \times 9$	-
	文献[54]	2013	Attenuation mask	$480 \times 270 \times 5 \times 5$	-
微透镜阵列	文献[26]	2005	MLA	$292 \times 292 \times 14 \times 14$	16 ms
	文献[56]	2014	MLA	$625 \times 434 \times 15 \times 15$	$1/(4\,000)$ s
	文献[58]	2006	MLA	$120 \times 120 \times 17 \times 17$	$1/15$ s
	文献[28]	-	Multi-Focus MLA	Depends on type	-

像传感器通过多次曝光捕获不同视角的多个光场样本,典型的方法是使用安装在机械机架上的传感器来测量不同位置的光场,如图3(b)所示. 斯坦福大学的计算图形实验室分别利用机械臂和 Lego Mindstorms 机械臂^[50]按照设计好的路径和位置操作一个传统相机捕获光场. Kim 等^[51]按照相同的思路设计了一组仅有一个方向自由度的光场相机,用于室外环境的光场捕捉. Liang 等^[52]利用可编码的光圈通过多次曝光来捕捉光场. 通过将光圈编码到不透明的纸缝或者液晶阵列中,允许相机记录不同视角的光线亮度,多次曝光并合成不同视角记录的光线信息获得场景的光场.

基于掩模的光场相机通过在主透镜和传感器之间插入可编码的掩模来捕捉光场,如图3(c)所示. Veeraghavan 等^[53]将这种方法称为“Dappled Photography”,相比于直接记录各个方向的光线,利用掩模对进入相机的光线进行光学调制,后通过傅里叶变换进行光场的重构. 得益于学习技术的发展,Marwah 等^[54]提出一种基于衰减掩膜的光场相机,利用超完备字典技术从二维投影影像中恢复出光场.

基于微透镜阵列的光场相机是在主透镜和传感器之间放置一个微透镜阵列,将经主透镜后各个方向汇聚的光线重新打散,达到记录四维光场的目的. 此时,传感器记录的是通过微透镜的单个方向光线的强度,而不是所有方向光线强度的总和. Adelson 和 Wang^[55]于1992年设计了一种新型的单镜头光场相机,可以捕捉 5×5 视角的影像. Ng 等^[26]通过将一个 296×296 的微透镜阵列安装在主镜头和传感器之间,设计了一种手持式光场相机,可以捕捉 11×11 个视角的影像. 如图4(a)所示,标准型光场相机每一个微透镜位置记录一个空间点信息,而且微透镜的焦距 f_M 等于微透镜阵列和传感器之间的距离 B ,商业相机 Lytro 和 Lytro Illum 均属于这个范畴^[56]. 通过调整微透镜阵列与主透镜和传感器的位置,Lumsdaine 和 Georgiev^[57]提出了聚焦型光场相机. 图4给出了不同模式下聚焦型光场相机主透镜、微透镜、传感器之间的距离关系. 如图4(b)所示,如果微透镜阵列安装在主透镜焦平面之前,即 $f_M < B$,这种设计被称为开普勒模式;如图4(c)所示,如果微透镜阵列安装在主透镜焦平面后方,即 $f_M > B$,这种设计被称为伽利略模式. 目前,聚焦型的光场相机设计已经被应用到工业光场相机 Raytrix^[28].

总的来说,多传感器光场相机不需要像基于微透镜阵列的光场相机的特殊光学元器件,但是在同步拍摄、相机标定和图像畸变校正方面存在较大的挑战,而且相机阵列的位置受限于相机的大小. 基于时间序列光场相机的方向分辨率则更多是受限于控制它的机械臂的精度,理论上可获得更高的方向分辨率. 除此之

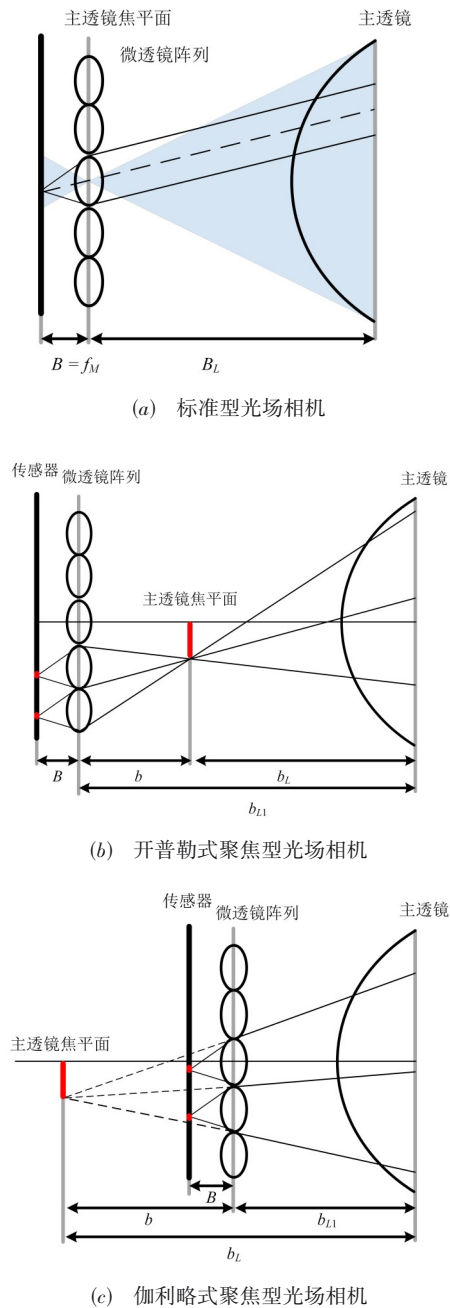


图4 基于微透镜阵列光场相机内部结构示意图

外,只需对一个相机进行标定. 但是基于时间序列的光场相机对装置的精密性要求较高,而且光场数据不是一次拍摄获得,因此更适用于静态场景. 对于基于掩膜的光场相机,由于掩膜的透光率无法达到100%,该方式获取的光场数据信噪比较低,目前研究较少. 基于微透镜阵列的光场相机体积较小,且能保证同时成像,但利用二维的影像同时记录四维的空间和方向信息,导致空间和角度维度的分辨率权衡问题. 除此以外,研究人员仍在研制新型光场相机,比如在 *Nature Photonics* 刊登了 Shehzad 等^[59]提出的石墨烯光场相机,在主镜头

后利用多层以石墨烯为材料的传感器来接收不同深度物体的光线,在不损失空间分辨率的情况下,记录了光场.

3 光场相机标定

相机标定指的是建立空间中一点与其在影像中对应像素之间的几何关系.不同类型光场相机内部构造不同,一般情况下需采用不同的标定方法.多传感器光场相机和基于时间序列的光场相机可被视为特殊的多视相机阵列,因此可采用传统相机阵列的标定方法.在传统的多视几何标定方法中,多个在不同姿态下的相机被定义为一组无约束的光线,被称为广义相机模型(Generalized Camera Model, GCM).Pless等^[60]对利用GCM标定时存在的问题进行了详细的讨论.因此,本文主要关注基于MLA的光场相机的标定.

基于MLA的光场相机因在传统相机的传感器前方添加了微透镜阵列,导致其投影模型和畸变模型有别于传统相机.而且,光场影像中空间信息和方向信息紧密地耦合在一起,一个物方点会在传感器上多次成像.因此,传统标定算法不适用于基于MLA的光场相机,标定的难度主要在于恢复光线在相机内部的传播路径,包括构建相机投影模型和构建相机畸变模型,并找到光场影像中像素与光线的对应关系.目前,大多将光场相机的主镜头和微透镜阵列的成像模型分别看作薄透镜模型和针孔模型,以此构建完整投影模型,忽略MLA的畸变,在此基础上,采用二维的棋盘格标定板作为物方标定物来确定物方和像方的对应关系.

除此之外,基于MLA的光场相机记录的光场影像是按一个个微透镜图像排列的,由于微透镜之间的距离不固定且会存在旋转和平移,导致微透镜图像排列不规则.而且,微透镜阵列是按六边形进行排列,以增加传感器的利用率,因此得到的每个微透镜图像也为六边形.为便于后续标定处理,部分方法首先将原始光场影像解码为一组等基线的传统多视影像,并在此基础上进行相机标定;另一类方法则直接在原始光场影像上进行标定.依据标定时是否采用原始光场影像,标定方法可大致分为两类:基于解码后光场影像的标定方法和基于原始光场影像的标定方法.

3.1 基于解码后光场影像的标定方法

为简便地处理原始光场影像,可首先将影像数据解码为一组不同视角的传统二维的子孔径图像.目前常用解码的方法有2种.

一种是Dansereau等^[61]提出的基于微透镜下覆盖像素数目和相邻微透镜之间的几何关系进行解码,进而提取子孔径图像.如图5(a)所示,原始光场影像进行去除马赛克和利用白图像消除衍射影像后,得到的光

场数据微透镜图像之间的间距不是整数,且存在旋转.首先对图像进行重采样、旋转和缩放操作使所有微透镜图像中心能够位于像素的中心,结果如图5(b)所示.下一步,根据每个微透镜下的像素数目,把所有像素归类为一个包含相同像素数、重叠的微透镜图像,如图5(c)和图5(d)所示.其中, (k,l) 表示微透镜在微透镜阵列中的坐标位置, (i,j) 表示每个微透镜图像内部像素的坐标.上述操作的像素都是在六边形的像素上完成.图5(e)和图5(f)则表示把六边形的像素转化为传统的方形像素以及把长方形像素进一步重采样为正方形像素的过程.至此,得到了解码后的光场数据 $L(i,j,k,l)$.

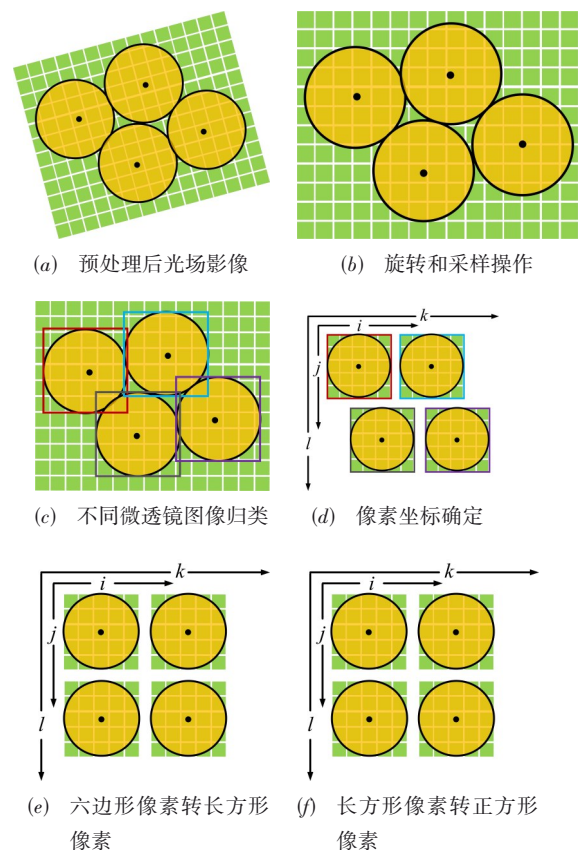
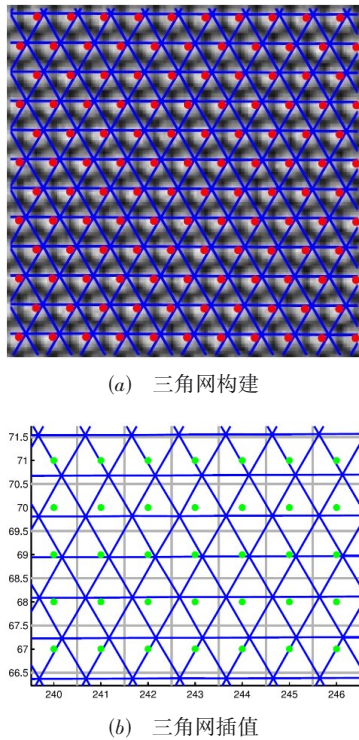


图5 基于微透镜几何关系的原始光场影像解码过程

另一种是Bok等^[62]提出的利用三角构网的方法提取子孔径图像.如图6(a)所示,每个微透镜下距离中心位置相同距离点连接起来构成的三角网,其中红点为每个微透镜图像中心;图6(b)中绿点为目标点像素,每一个绿点都位于对应三角网内部,根据三个角点像素值进行三角网插值可得其像素值.无论哪种解码方法,都需事先找到每个微透镜图像的中心像素,目前通用的做法是利用光场相机的白图像.由于衍射现象,在白图像中,每个微透镜图像中最亮的点就是微透镜的中心.

图6 基于三角构网方法的原始光场影像解码^[62]

在完成原始光场影像解码的基础上, Dansereau 等^[61]提出一种包含 15 个参数的标准型光场相机标定方法. 通过构建每个像素与对应光线之间的几何关系, 并采用分步求解的方法, 完成了标定参数的求解. 分步求解过程中, 首先基于解码得到的子孔径图像利用传统标定方法得到外方位元素; 然后在不考虑畸变的情况下, 采用一种线性的方法进行求解, 得到所有参数的初值; 最后在考虑畸变的情况下, 对初始评估结果进行非线性优化. 但是求解出的标定参数没有特定的物理现实意义. Zhang 等^[63]进一步发现 Dansereau 等^[61]构建的影像模型中存在参数冗余, 提出一种仅包含 6 个参数的多投影中心标定方法 MPC (Multi-Projection-Center). 类比传统 GCM 方法, 他们推演了光场影像解码出的四维光场与物方三维结构的几何投影关系, 实现相机内部和物方世界的 MPC 坐标的线性转换. 此外, 针对该投影模型, 他们还提出一种内外方位元素的初始估计方法, 通过最小化重投影误差进行非线性优化. 与 Dansereau 等^[61]的方法相比, MPC 具有更少的参数冗余, 可同时应用于标准型和聚焦型光场相机.

不同于直接构建物方空间点和像方角点的投影模型, 研究人员也通过深度和核线平面影像 (Epipolar Plane Image, EPI) 等信息间接完成光场相机的标定. Ji 等^[64]利用光束法平差确定子孔径影像所对应视角在主透镜上的位置, 完成了光场相机标定. 他们首先找到中心子孔径图像在主透镜上的位置作为参考, 其次基于

主透镜的焦距估计空间点在不同子孔径图像下得到虚拟像点, 最后利用光束法平差求解子孔径图像视角在主透镜上的位置, 但方法的精度受限于方向分辨率的大小. Pertuz 等^[65]通过构建深度与重聚焦参数之间的对应关系完成了光场相机的标定. 对应关系的建立是基于对光场相机成像过程的一阶光学分析推导得到. 针对建立的几何模型, 提出了一种模型参数求解方法. 该方法无需了解光场相机的内部参数, 仅需利用传统的棋盘格影像, 但该方法只适用于标准型光场相机. Zhou 等^[66]通过推演 EPI 中线的方向和截距与其对应空间点的三维坐标的投影关系, 提出了一种基于核线空间的 4 参数几何标定模型. 具体地, EPI 中线的方向对应于物方点的深度, 水平和垂直方向的截距分别对应于物方点的 x 和 y 坐标. 该方法利用了更少的参数, 但其精度依赖 EPI 中线的检测精度.

虽然以上方法可以取得相对较好的标定精度, 但无论是直接法还是间接法, 均未考虑微透镜的畸变, 造成一定的精度损失. 而且, 构建的投影模型相对简单, 没有考虑具有不同 MLA 结构的光场相机, 比如多焦点 MLA 光场相机包含几种不同排列的微透镜阵列, 因此针对不同类型光场相机构建需设计更全面的投影模型以及畸变模型.

3.2 基于原始光场影像的标定方法

如图 7(a) 和图 7(b) 所示, 对于棋盘格标定板的原始光场影像, 根据每个微透镜图像像素的组成, 可将标定板的微透镜图像分为 3 类: 包含线特征的微透镜图像、包含点特征的微透镜图像和背景类微透镜图像. 一些研究人员直接利用原始光场影像的点、线特征来完成相机标定. 在几何校正后从原始影像中提取的子孔径图像的质量要优于在解码后光场图像上按照等间距提取的子孔径图像^[62].

由于微透镜图像较小, 比如 lytro Illum 为 15×15 像素, 传统的角点检测算法应用于微透镜图像时, 精度较差. 为此, Liu 等^[67,68]提出一种基于原始光场影像角点特征的标定方法. 首先, 采用 2 组模板影像, 若待检测点周围的像素分布与模板影像一致, 则判定为角点特征, 如图 7(c) 所示. 在此基础上, 他们将光场相机的固有参数分为与传统相机一致的内参数和光场相机特有的内参数两类, 首先对这两部分内参数进行分步求解, 再通过最小化重投影误差对初始估计做进一步优化.

如图 8 所示, 当相机与标定板存在较大的姿态变化时, 出现大量错误检测和漏检的情况. 除此之外, 由于优化方法对初始估计不敏感, 分步求解难以提高最终精度. 针对以上问题, Liu 等^[69]提出一种新的基于角点特征的标定方法. 首先在原始图像上找到包含角点特征的区域, 并提出一种圆形边界策略在这些微透镜图

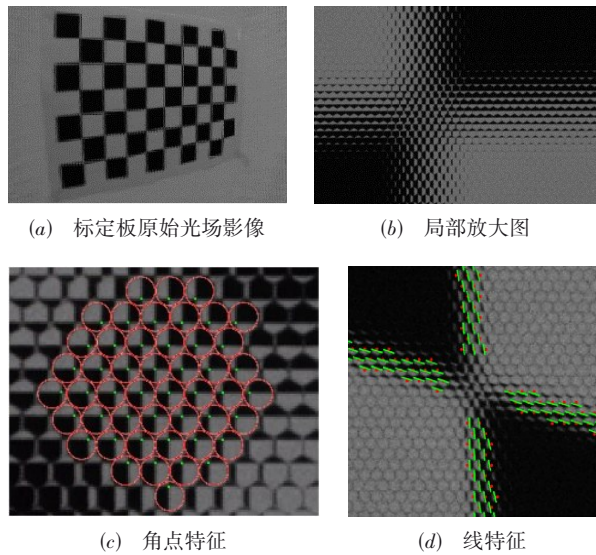
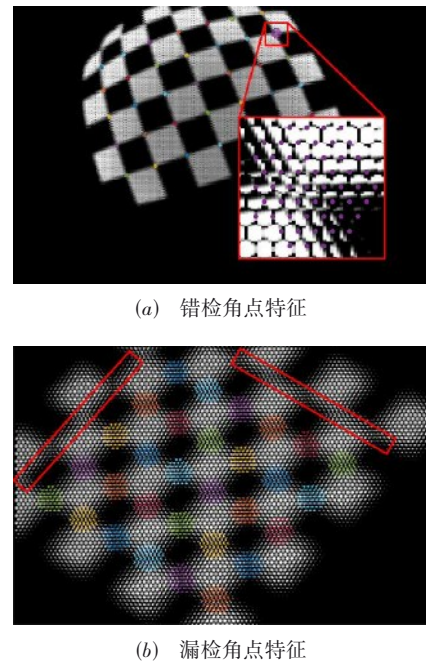


图7 光场相机拍摄的原始标定板影像以及提取的点、线特征

像上检测角点特征;然后基于中心子孔径图像,利用常规标定方法计算初始外参数,并根据角点及其对应中心的几何关系,采用线性方法计算初始内参数;最后,所有的初始估计都通过非线性优化得到进一步的细化.由于提高了角点的检测精度,Liu等^[69]的方法获得了更好的标定结果.针对多焦点聚焦型光场相机标定,Nousias等^[70]同样提出一种基于角点特征的方法.首先,基于聚焦程度区分不同类型微透镜,并将其按照标定板上角点特征所对应点顺序进行排列;其次,通过寻找包含角点特征的微透镜影像中近似沙漏形状两条线的交点确定角点特征的位置;最后,构建空间点与对应角点特征的几何关系,实现多焦距光场相机的标定.

考虑到微透镜影像中包含大量的线性特征,Bok等^[62]利用从原始光场图像上直接提取的线特征,提出一种基于原始光场影像线特征的标定方法,实现了光场相机标定.该方法从微透镜图像中利用模板匹配提取直线特征,如图7(d)所示,构建物方与像方的几何投影关系并进行标定参数求解.如图9所示,从一个物方点 P_c 射出的光线首先穿过主透镜后汇聚于一个虚拟像点 P_v , P_v 通过微透镜后投射到传感器上,得到 P_c 在一个微透镜(中心为 p_c)下的像点 $p(x,y)$;其次,利用得到的几何投影关系计算出初始的内方位参数和外方位参数;最后,通过最小化所提取直线上离微透镜中心最近一点对应的光线与空间中直线之间的距离来进行非线性优化.

由于原始光场影像上的线特征较多,算法具有较好的稳定性.但在投影过程中由于非线性项的存在,空间中的直线投影到微透镜影像上时并不是一条严格的直线,因此非线性优化时所选取点对应的光线并不会

图8 标定板光场影像角点特征检测失败示例^[67]

与空间中的线相交,这会直接影响到标定的精度.此外,微透镜影像中线的检测精度也受限于线检测过程中所选取模板个数的多少.未来可考虑利用点、线特征结合的方法,首先利用大量的线特征提供初值,然后利用角点特征在非线性优化过程中具有的优势,实现光场相机的准确标定.

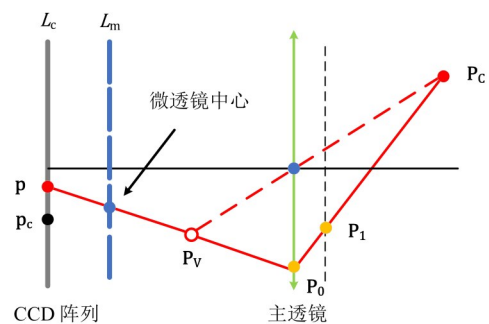


图9 标准光场相机投影模型

3.3 典型光场标定方法对比实验分析

本文选定了5种具有代表性的光场相机标定算法,包括:先解码后标定的方法 Dansereau等^[61](记作DPW),MPC^[63],以及Bok等^[62]、Liu等^[67]和Liu等^[69]提出的基于原始光场影像的标定方法,分别记作BJW,SCC和LFC.实验数据选用了Dansereau等^[61]提供的5组标准测试数据A、B、C、D和E(<http://dgd.vision/Tools/LFToolbox/>),每组数据中包含多张不同姿态、不同位置拍摄的棋盘格标定板影像.格网的大小也不全相同,其中数据集A和B、C和D、E的格网大小分别为

3.61 mm、7.22 mm、35.1 mm. 由于原始光场影像上的线特征和点特征只有在相机拍摄失焦的情况下才存在,因此5组数据中只有部分数据可以用于评估基于原始光场影像的标定方法. 具体地,A组数据中第1、2、4、5、6幅共5张影像,B组数据中第1~5、15~18幅共9张影像,D中第1~15幅共15张影像,数据集C和D中影像均在聚焦情况下拍摄,因此均不能使用. 每幅图像的大小为3 280×3 280像素,微透镜图像的半径为5像素,考虑到位于边缘视角的子孔径图像会存在比较大的缺陷,仅评价位于内部视角7×7子孔径图像的像方的投影误差(Re-Projection Error, RPE)和物方投影误差(Ray

Re-Projection Error, R_RPE). 具体实验过程中,DPW、BJW、SCC和LFC利用了原作者提供的最新开源代码,而MPC则直接利用了作者原文中提供的结果.

表2给出了5种光场标定方法的对比实验结果. 从结果看出,在数据集A和B上,DPW和MPC取得了显著优于BJW、SCC和LFC的结果,而在数据集E上,BJW与DPW和MPC的精度在同一水平,LFC的重投影误差最小. 这可能是因为数据集A和B中,基于原始光场影像的标定方法能够利用的有效影像数目明显少于先解码后标定的方法,而在数据集E上,两类方法能够利用的影像数目相差不多,因此两类标定方法的精度相差不多.

表2 几种典型光场标定算法在Dansereau等^[61]提供的5组测试数据集上的标定结果

数据集/ 网格大小(毫米)	评价标准	先解码后标定的方法		基于原始光场影像的标定方法		
		DPW ^[61]	MPC ^[63]	BJW ^[62]	SCC ^[67]	LFC ^[69]
A/3.61	PE(像素)	0.228(18)	0.220(18)	0.730(5)	-	0.713(5)
	R-PE(毫米)	0.082(18)	0.081(18)	0.218(5)	-	0.214(5)
B/3.61	PE(像素)	0.158(18)	0.157(18)	0.483(9)	0.346(9)	0.464(9)
	R-PE(毫米)	0.059(18)	0.057(18)	0.147(9)	0.097(9)	0.142(9)
C/7.22	PE(像素)	0.195(18)	0.175(18)	-	-	-
	R-PE(毫米)	0.130(18)	0.112(18)	-	-	-
D/7.22	PE(像素)	0.167(18)	0.148(18)	-	-	-
	R-PE(毫米)	0.115(18)	0.105(18)	-	-	-
E/35.1	PE(像素)	0.336(18)	0.273(18)	0.284(15)	0.686(15)	0.224(15)
	R-PE(毫米)	0.384(18)	0.539(18)	0.558(15)	0.994(15)	0.348(15)

注:标定结果后括号内的数字为实验过程中所利用的影像数目

同类方法内,MPC总体上取得了略优于DPW的结果,而且由于MPC的标定参数个数不到DPW的一半,因此MPC的标定速度更快. 相较于BJW和LFC,SSC表现不稳定,在数据集B上取得了优于BJW和LFC的结果,但在数据集E上大幅差于BJW和LFC,甚至在数据集A上标定失败. 这是由于SSC的角点检测方法鲁棒性较差. 当标定板所在的世界坐标系与相机坐标系之间的姿态变换较大时,SSC角点检测失败,进而导致标定失败,在数据集A上就是这种情况,这也反衬出同样基于原始光场影像角点特征的LFC的角点检测算法对不同影像姿态具有较好的抗性. 除此之外,LFC的总体表现优于BJW,这是因为最终的优化结果主要取决于非线性优化过程,而空间中的直线投影于二维影像上的过程是非线性的,因此影像上得到的并非严格意义的直线,只是因为微透镜影像较小,可近似看作直线,因此BJW的表现会差于LFC. 需要注意的是,较小的残差输出仅表明在自身算法的优化过程中有较好的收敛,但也可能出现过拟合现象. 而且,不同算法的标定参数不同,尤其是基于原始光场影像的直线或者角点特征本身以及数目不同,也会影响到残差计算过程中的公平性. 因此,本节仅是对不同光场相机标定方法的

初步评估,若要更全面地评估不同方法,需构建更多包含不同条件的数据集,并利用标定的相机参数计算影像深度,并与真实值进行比较.

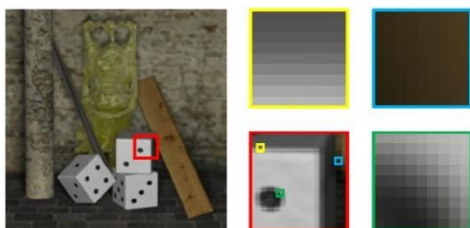
4 光场影像三维深度信息恢复

光场固有地包含多视角信息,可据此估计场景的深度,但光场结构的特殊性给三维深度信息的恢复带来了挑战,研究人员在这方面做了大量工作. 根据光线的不同排列方式,光场有3种可视化表达:多视子孔径图像,角度域影像^[71]和EPI^[72]. 具体地,如果将通过主透镜上同一位置、不同空间位置的光线按空间顺序排列在一起得到一个视角的子孔径图像,而不同视角子孔径图像排列在一起,组成多视角子孔径图像,如图10(a)所示;如果将一个空间点各个视角光线按视角顺序排列在一起,可以得到该空间点的角度域影像,如图10(b)所示;在多视角影像中,如果沿某个方向选择多视角影像并进行前后排列,并沿相同方向取切面,则该切面影像即EPI,如图10(c)所示. 基于不同表达形式,当前光场三维深度信息恢复方法分为基于多视匹配^[73-75]、基于角度域影像^[76-82]和基于EPI^[83-89]的光场三维深度恢复. 此外,得益于深度学习技术的发展,研究人员也

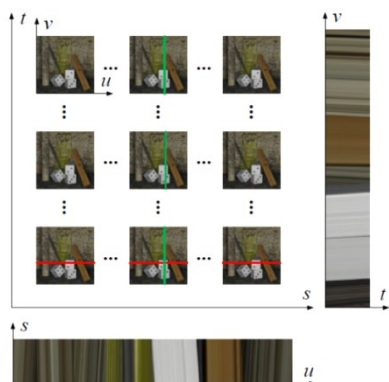
提出诸多基于光场不同可视化表达的利用深度学习进行光场三维信息恢复的方法^[90-93].



(a) 多视子孔径影像



(b) 角度域影像



(c) 核线平面影像(EPI)

图 10 光场的三种可视化表达^[40]

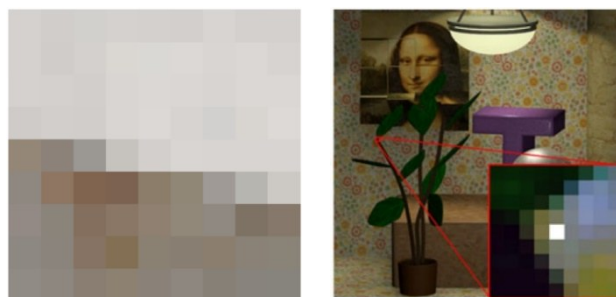
4.1 基于多视影像的光场三维深度恢复

针对多视角子孔径图像,可利用传统多视的方法来恢复场景的三维信息,但由于不同子孔径图像之间的基线较短,传统的多视角匹配算法性能不佳.为此,Heber等^[73]指出从光场数据提取的子孔径图像与在主动波前采样设置中捕获的图像相似,可以通过测量子孔径图像上图像点旋转密集场的变分模型来计算深度.Yu等^[74]提出一种线辅助的光场三角构网和立体匹配算法.首先将双线性子空间映射到线约束,然后再应用到有约束的Delaunay三角构网,减少了构网的误差.在此基础上,提出一种线辅助的图割算法,不同视角子孔径图像的线约束下完成匹配,提高了立体匹配的可靠性,但是该算法没有考虑遮挡存在的情况.Jeon等^[75]基于多视的思路提出一种光场深度估计算法.首先利

用频域的亚像素偏移,设计了绝对差和梯度差两种数据代价,提高了窄基线子孔径图像的匹配精度,并进一步采用多标签优化模型和迭代优化模型对初始深度图进行优化,但是该方法存在边缘过度平滑的问题.

4.2 基于角度域影像块的光场三维深度恢复

通过重聚焦原始光场图像,可得到一系列具有不同焦距的影像,当聚焦到正确的深度时,角度域影像内的像素会显示出一致性.Tao等^[76]提出利用散焦和像素一致性相结合的方法计算深度.他们把角度域影像内像素的方差值作为像素一致性线索,并把重聚焦图像的梯度值作为散焦的线索,然后基于置信度的马尔可夫随机场模型将这两种线索结合起来估计三维深度.在此研究的基础上,Tao等^[77]又增加了阴影信息作为正则化,并利用中心孔径图像改进了原始的散焦和像素一致性线索.具体地,以重聚焦图像和中心视角图像的平均强度差作为散焦线索,以包含不同视角的像素块与中心视角像素的差值之和作为像素一致性线索,得到更为准确的深度信息.在未遮挡区域,上述方法取得了比较好的结果,但对于被遮挡像素,如图11所示,角度域影像内不同视角像素的一致性被破坏,引起估计误差.



(a) 仅有一个遮挡物^[79] (b) 存在多个遮挡物^[81]

图 11 在遮挡情况下角度域影像的像素组成

Wang等^[78,79]研究发现,只存在单一遮挡物时,在遮挡区域分割被遮挡视角和未被遮挡视角的直线与空间域中的遮挡边缘具有相同的方向,基于此提出了遮挡自适应的深度估计算法.首先设计了三种遮挡像素检测算子找到被遮挡像素,进而对被遮挡和未被遮挡像素分开处理.具体地,在遮挡区域,仅利用一部分未被遮挡的视角评估深度,提高被遮挡像素的深度评估精度,但该方法仅能应用于存在一个遮挡物的情况.Chen等^[80]提出一种基于颜色和空间的双边一致性度量(Bilateral Consistency Metric, BCM),提高了在遮挡区域深度估计的可靠性.BCM计算了角度域影像中每个视角被遮挡的可能性大小,以此找到不同深度下未被遮挡的视角,并计算代价量.基于不同深度下代价量的变化曲线,确定深度值的可靠性,提高了深度估计的鲁棒

性. BCM虽然在遮挡区域表现较好,但不能应用于方向分辨率较低(比如 5×5 以下)的光场影像,且对噪声比较敏感. 针对存在遮挡和噪声的复杂场景,Williem等^[81]提出了角熵(Angular Entropy, AE)和自适应散焦(Adaptive Defocus, AD)两种数据代价. AE统计大部分像素亮度的概率,计算不同角度域影像的信息熵,熵值时对应的深度为正确深度. AD仅测量整个角度域影像中一部分子区域的散焦响应,增强了对遮挡的抵抗性. 除此之外,AE和AD都添加了额外的颜色相似性约束,获得了较好的深度估计效果. 在AE和AD的基础上,Williem等^[31]进一步提出了约束角熵代价(Constrained Angular Entropy, CAE)和约束自适应散焦代价(Constrained Adaptive Defocus, CAD). CAE基于颜色相似性对每个像素进行加权,而不是赋予所有像素相同的权重,这使得遮挡像素的贡献更小,更好地抑制了噪声. CAD没有在选定的子区域上添加颜色约束,而是将该约束放置在所有子区域并计算最小代价,降低了遮挡的影响. 最后,利用马尔科夫随机场模型(Markov Random Field, MRF)对两种度量进行综合,进一步提高了对遮挡和噪声的鲁棒性,从而提高了在复杂场景下的估计精度. 图12给出了BCM和综合利用AE和AD方法在有无噪声情况下恢复得到的深度结果图对比. 可以看出,噪声对BCM方法的影响较大,而后者对噪声具有较好的抵抗性.

4.3 基于EPI的光场三维深度恢复

EPI中线的方向对应于图像的深度,可通过分析线结构的方向估计深度. Wanner等^[82]使用结构张量计算EPI中线的方向并评估其可靠性,然后使用变分法优化初始深度值,得到亚像素级的精度. 该方法从一定的深度范围内直接寻找正确的深度值,而不用将深度范围离散化处理,可得到连续的深度图,而且速度较快,但该方法容易受到噪声和遮挡的影响. Li等^[83]在Wanner等^[82]的基础上,增加了颜色一致性约束,提高了结构张量在EPI上深度估计的可靠性. 基于初始估计深度对整幅图像进行分割,并采用稀疏线性系统,在不同区域使用两个不同的相似性矩阵,对初始评估结果进行优化,提高了精度和效率,但依然没有考虑遮挡. Zhang等^[84]使用EPI中线上像素的强度、梯度和空间一致性来找到最佳角度,并利用双线性内插提高初始评估角度的连续性. 在此基础上,计算每个像素的置信度,找到评估结果不可靠的像素,基于图像本身的结构,将周围可靠像素的深度扩展到不可靠像素,完成整幅图像的深度评估,但是该方法仅能应用于具有较大方向分辨率的光场影像. Suzuki等^[85]针对大视差范围光场数据,提出对EPI进行斜切处理来提高评估精度. 他们发现,结构张量在EPI中线的方向接近垂直时评估精度较高,

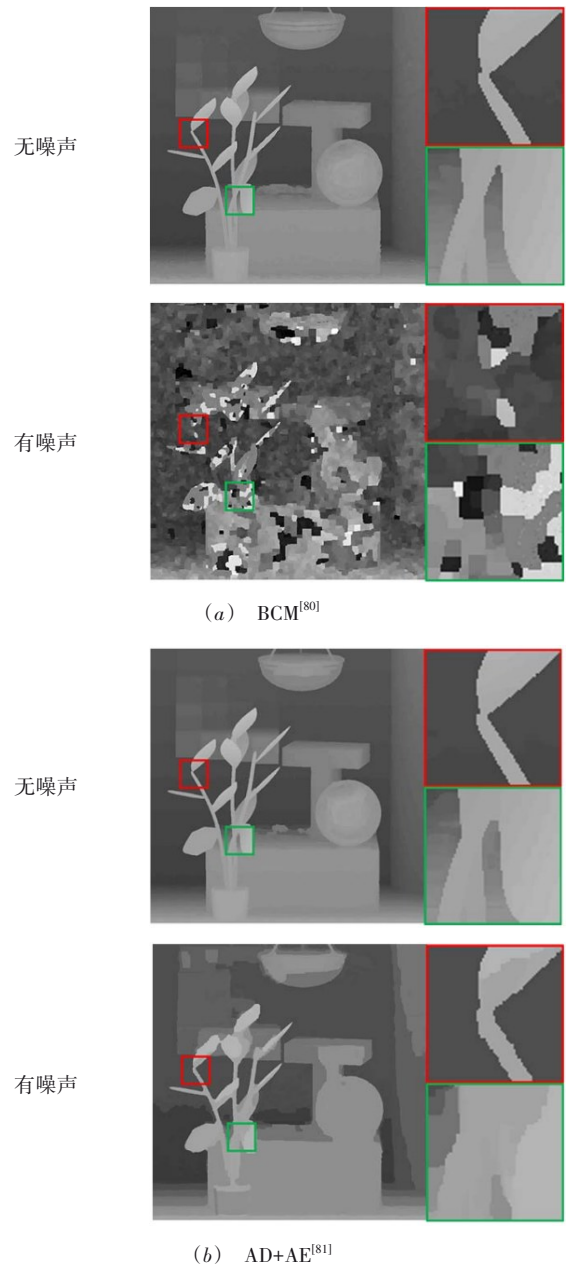


图12 在无噪声和有噪声情况下深度估计结果对比

首先对EPI进行不同程度的斜切处理并计算置信度,然后通过滤波将高置信度的结果集成到最终的视差图中,提高了在大视差情况下的评估精度,但该算法未考虑遮挡. Zhang等^[86]提出了一种基于EPI的旋转平行四边形算法SPO(Spinning Parallelogram Operator). SPO首先通过计算EPI中平行于目标线两侧平行四边形区域的最大加权直方图距离来寻找单方向EPI上的最优方向;然后提出一种可靠的深度置信度评估算法,对水平和垂直方向的评估结果进行聚合;最后利用边缘保持滤波算法进一步优化评估结果. 虽然SPO没有对遮挡和噪声进行明确的处理,但对遮挡和噪声都有较好的抵抗性.

图 13 是一张国际棋盘格光场影像^[80]. 场景中存在着复杂的遮挡关系, 该种情况下提取的 EPI 中线结构被破坏(如图中红色和黄色框选区域), 导致无法正确提取影像的深度.

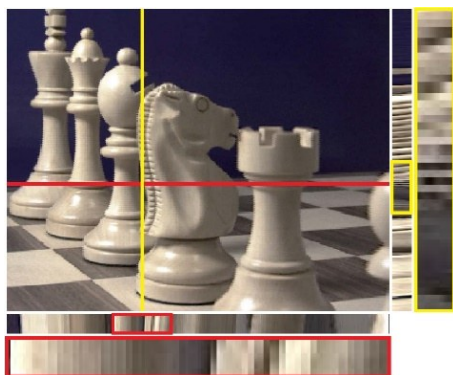


图 13 EPI 在遮挡下的表现^[80]

Sheng 等^[87]在 SPO^[86]的基础上, 利用多方向的 EPI 上的有用信息, 提高了在遮挡区域的深度估计精度. 首先提取多方向的 EPI 并评估深度, 然后利用多次评估的深度值计算两个遮挡线索, 找出被遮挡像素. 对处于遮挡边缘的像素点, 仅利用与遮挡边界方向相同的 EPI 的评估结果, 减少了遮挡的影响. Schilling 等^[88]提出一种综合利用深度和遮挡的深度评估模型. 具体地, 在评估过程中, 动态更新已评估像素的深度, 以此对待评估像素的各视角信息进行判断, 找出未被遮挡视角, 计算代价量. 在此基础上, 利用周围信息进行局部优化. 虽然该算法并未进行全局优化, 但是在边缘和弱纹理区域均表现较好.

4.4 基于深度学习的光场三维深度恢复

深度学习算法已经在图像处理、目标检测、自动驾驶等领域取得了前所未有的成果. 针对光场影像三维深度信息恢复这一问题, 研究人员采用深度学习领域前沿算法开展研究, 探索深度学习技术在光场三维重建中的应用潜力. Shin 等^[89]提出一种基于全卷积神经网络的光场深度估计方法 EPINET. 首先, 利用光场几何特性设计一种端到端的神经网络结构; 其次, 构建水平、垂直以及左右对角线多流网络; 最后, 使用 16 个光场合成图像, 对光场特定数据进行增强, 并从这 16 个合成光场图像中随机抽取 23×23 的灰度图像块进行训练. 该算法不仅在合成光场数据的各项指标上均表现较好, 而且算法的可靠性在一小部分真实光场数据上也得到了验证. 为了更充分利用光场影像中有效的多视角信息以及剔除无用信息, Tsai 等^[90]提出一种基于新的神经卷积神经网络光场深度估计算法. 首先通过构建视图选择模块生成一幅注意力图, 找出每个视角信息的潜在重要性大小, 然后利用光场不同视角信息的对

称性增强注意力图的对称性, 提高了深度估计的准确性, 获得了较好的结果. 不同尺度和方向的 EPI 可能会影响最终的深度估计精度. 为此, Zhou 等^[91]设计了一种基于 EPI 的尺度和方向自感知学习方法来估计深度. 具体地, 以每个像素的多向 EPI 块为输入, 设计了两种自适应尺度选择和方向融合的网络结构, 一种是尺度自感知结构, 另一种是用于融合多方向特征的共享权重网络, 提高了深度估计的准确性. Wang 等^[92]针对光场深度估计, 提出一种基于对极几何和图像分割的卷积神经网络. 首先利用对极几何来估计初始视差图, 然后将多方向 EPI 作为输入数据, 根据不同方向 EPI 的视差选择卷积块, 同时利用图像分割结果获得中心视差子孔径图像的边缘信息, 通过将两个部分的输出串联起来, 得到最终的深度图.

4.5 不同方法对比与分析

不同的数据形式在深度估计中有其自身的优势. 传统多视影像可以方便地利用周围信息, 但由于影像间基线很短, 易出现误匹配, 造成错误深度估计. 角度域影像能有效地利用光场重聚焦的特点, 且易于去除遮挡的影响, 但难以利用周围信息. EPI 将视差固定在一个维度, 利于提高深度估计的准确性, 但当前基于 EPI 的方法仅利用了线上或者线周围一方面信息, 降低了深度估计精度. 基于深度学习的方法在部分合成数据上取得了较好的结果, 但训练深度神经网络需要非常大的数据集, 目前还只有很少一部分包含深度真值的真实光场数据可用于学习, 因此基于深度学习的方法还难以应用于真实光场数据的深度估计.

为评估不同方法的性能, 本文选用了基于多视的 Jeon 等^[75]提出的方法(记作 MVP), 基于角度域图像的 Wang 等^[78]提出的方法(记作 OV)和 CAE^[31]以及基于 EPI 的 SPO^[86]4 种典型方法进行对比实验分析, 其中 CAE 和 OV 对遮挡进行了针对性的处理. 算法的实现利用了原作者提供的代码, 参数按照作者的设置. 实验数据利用了带有真实深度值的合成光场数据和真实光场数据. 合成数据是由 Wanner 等^[93]提供的光场数据深度估计算法标准测试数据集中的“Buddha”“Mona”“Stilllife”和“Papillon”数据, 便于进行定量评价; 真实数据采用了瑞士洛桑理工学院(École Polytechnique Fédérale de Lausanne, EPFL)提供的 Lytro Illum 相机拍摄的光场影像^[94].

对于所有的算法, 将最大和最小视差范围等分为 81 个深度标签以确保对于每种算法的每个深度标签的步长是一致的. 对于合成光场数据, 最大和最小视差的输入是根据真值确定的; 对于真实光场数据, 最大和最小视差的输入分别取 1 和 -1. 考虑到上述每种方法采用了不同的优化方法, 首先比较了未经优化的深度估

计结果(Local),以测试每个成本量的区分能力;然后在应用每种算法自身优化方法的情况下,对优化后的结果(Optimized)进行了比较.为了实现定量评估,采用100倍的均方根误差RMSE作为标准.

$$RMSE = \frac{1}{|N|} \sum_{p \in N} |GT(p) - d(p)|^2 \times 100 \quad (3)$$

其中, N 是待评估深度图的大小, GT 和 d 分别是真值和计算出来的深度.同时,对被遮挡区域单独评价,用 $RMSE_{occ}$ 表示其精度.被遮挡区域根据真值计算得到,具体地,如果真值深度图上某个像素在 x 和 y 方向上梯度值和的绝对值大于0.01,则认为该点是被遮挡.

表3分别给出了几种算法在优化前(Local)和优化后(Optimized)的定量对比实验结果,为方便查看,每组数据的最优结果用加粗标出.MVP在优化前和优化后

均表现不佳.OV整体表现相对平均,在遮挡区域则表现不稳定,在部分数据上取得了比较好的结果,比如在“Mona”数据优化后深度图的 $RMSE_{occ}$ 值为6.457,表现最好,但在“Buddha”数据上优化后深度图的 $RMSE_{occ}$ 值为13.238,表现最差.整体上,CAE在优化前表现最好,在四组数据上均取得了最好的结果,但其在优化后提升有限,仅从0.559提升到0.429;在遮挡区域,在“Buddha”“Mona”和“Stilllife”数据上则出现了优化后变差的情况,导致其平均 $RMSE_{occ}$ 从5.883变大为7.208.SPO在优化前均表现较差,但SPO在优化后有明显提升,比如SPO优化前RMSE和 $RMSE_{occ}$ 的均值分别为3.491和8.717,优化后分别提升到0.372和5.426,在整体和被遮挡区域精度均最高.

图14展示了几种算法在合成数据上评估的深度图

表3 几种典型光场深度估计算法在Wanner等^[93]提供的合成数据集上的定量实验结果

光场数据	类别	MVP ^[75]		OV ^[78]		CAE ^[31]		SPO ^[86]	
		RMSE	RMSE _{occ}	RMSE	RMSE _{occ}	RMSE	RMSE _{occ}	RMSE	RMSE _{occ}
Buddha	Local	4.388	12.135	0.769	7.412	0.434	7.323	3.947	11.647
	Optimized	1.147	12.380	0.990	13.238	0.628	11.626	0.386	8.398
Mona	Local	7.211	21.873	1.257	8.365	1.113	9.591	5.077	13.701
	Optimized	1.019	11.676	0.992	6.457	0.749	11.441	0.762	8.946
Stilllife	Local	0.676	2.661	0.430	2.088	0.147	1.741	2.690	2.531
	Optimized	0.533	2.249	0.403	1.947	0.124	1.782	0.135	1.566
Papillon	Local	4.301	12.173	0.846	4.517	0.540	4.878	2.250	6.987
	Optimized	2.086	8.539	0.288	2.977	0.216	3.982	0.205	2.793
平均值	Local	4.144	12.211	0.826	5.596	0.559	5.883	3.491	8.717
	Optimized	1.196	8.711	0.675	6.155	0.429	7.208	0.372	5.426

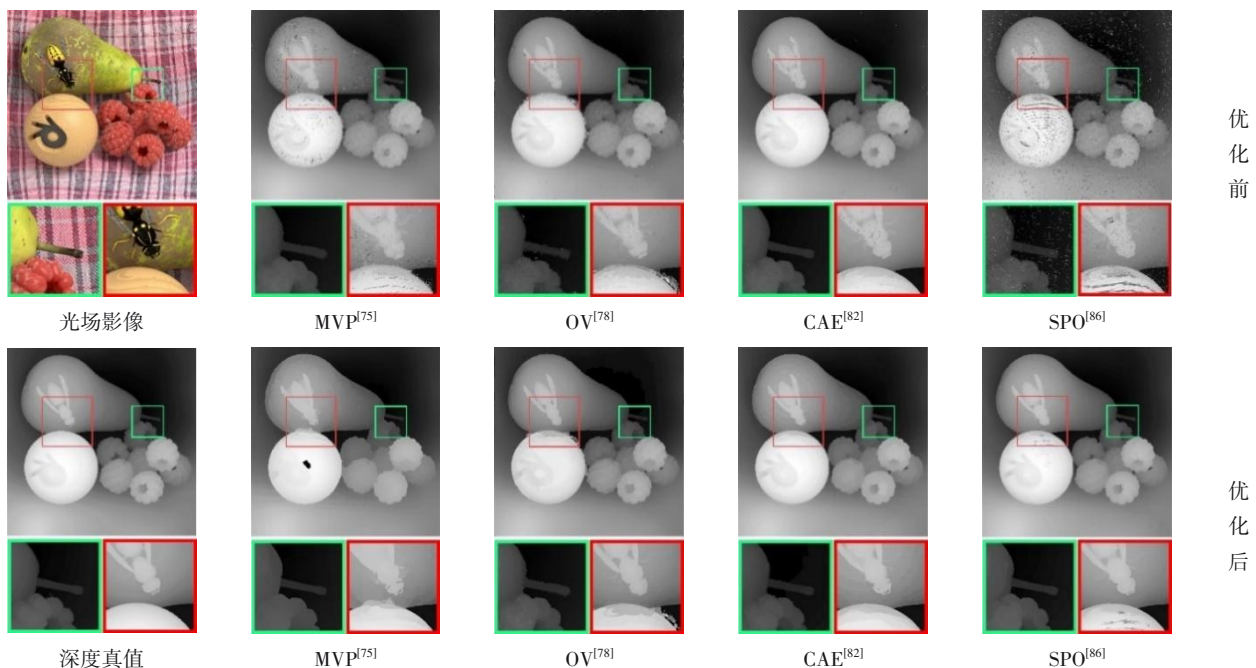


图14 合成光场数据上评估的深度图对比实验结果

对比实验结果.可以看出,MVP和SPO的深度图在优化前均存在大量被错误估计的深度噪点,这也是它们在定量结果上表现较差的主要原因.OV和CAE在优化前噪声较少,但OV在边缘出现了明显的错误,CAE则整体表现较好.优化后,MVP在平坦区域表现良好,但存在边缘过度平滑的情况;OV在表面和边缘恢复方面表现都比较平均;SPO和CAE在细纹理结构和纹理不连续区域都表现较好,同时在纹理匮乏区域均表现较好.

图15展示了几种算法在Lytro Illum拍摄影像上评估的初始深度图对比结果.数据中不仅包含具有挑战性的场景,例如遮挡、细结构和纹理匮乏区域,而且还

存在大量的噪声和失真的情况.从结果可以看出,MVP在优化前表现较好,但优化后存在过度平滑的问题,且在弱纹理区域出现大片连续区域的错误深度估计.OV受影像噪声影响较大,优化前存在较多噪声,优化后得到较大程度的提升.CAE在遮挡边界处表现良好,噪声较少,但在细结构和低纹理区域容易出错.SPO虽然在优化前的深度图中也存在较多噪点,但考虑到在合成数据中同样存在噪点,因此受光场影像质量影响不大;而且在优化后的深度图上,SPO在保持物体形状、恢复弱纹理表面和保持细结构方面表现最佳.

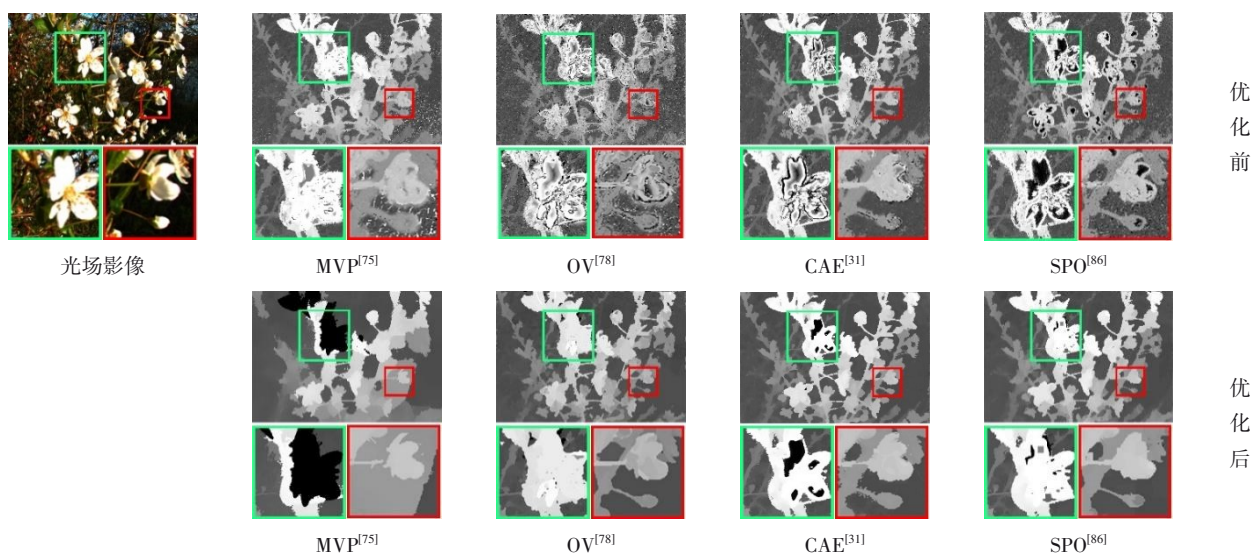


图15 真实光场数据上可视化对比实验结果

5 开放获取的光场数据和算法

大量的公开光场相机数据以及光场影像处理算法,促进了光场领域的发展.本文对目前的公开资源进行了总结.表4列出了目前开源的光场数据,包含合成光场影像、多传感器光场影像、基于时间序列的光场影像以及标准型/聚焦型光场影像.斯坦福大学计算图形实验室在光场研究领域非常活跃,先后公布了利用相机阵列、Lego gantry、gantry、光场显微镜、Lytro Illum拍摄的光场影像,用以深度估计和多次拍摄光场影像的重建性能测试^[50,95].麻省理工学院的Wetzstein^[96]、海德堡大学的Wanner等^[93]和Honauer等^[97]发布了包含真值的合成数据集,用于测试深度估计算法.洛桑理工学院的Rerabek等^[94]、罗马大学的Paudyal等^[98]和慕尼黑工业大学的Hazirbas等^[99]发布了包含不同场景类别和不同数目的Lytra Illum光场影像数据集.慕尼黑工业大学的Zeller等^[100]公布了利用双目相机和Raytrix R5同

步拍摄的数据,用以对比单目、双目和光场相机在三维重建上的不同表现.法国国家信息与自动化研究所(Institut National de Recherche en Informatique et en Automatique, INRIA)先后发布了Lytra Illum、Raytrix R8以及合成光场影像数据集^[101],用以测试深度估计算法、场景识别以及标准型光场相机和聚焦型光场相机的数据联合处理问题.瑞典中部大学的Ahmad等^[102]发布了同时包含Lytra Illum和Raytrix R29影像的数据集,用于测试标准型和聚焦型光场影像的匹配结果.英国的贝德福特大学的Hahne等^[103]公布了自主研发光场相机拍摄的光场数据.

表5列出了目前与光场三维重建相关的开源算法,包括光场相机的标定和光场影像三维深度恢复两个方面.其中大部分是国外机构开源,另外也有部分国内机构如海军航空大学、西北工业大学和台湾大学等,大大促进了该领域的发展.

表 4 公开的光场数据小结

数据集	获取方式	发布机构	时间
The new Stanford LF archive ^[50]	Multi-Camera Array; Lego gantry; Light field microscope	Stanford University	2008,2016, 2019
The Stanford Lytro LF archive ^[50]	Lytro Illum		
The Stanford Multiview LF ^[95]	Multi-view LF captured by Lytro Illum		
MIT LF archive ^[96]	Synthetic	Massachusetts Institute of Technology	2011,2012, 2013
4D LF dataset ^[93,97]	Synthetic	Heidelberg University	2013,2016
EPFL LF dataset ^[99]	Lytro Illum	École Polytechnique Fédérale de Lausanne	2016
SMART Dataset ^[100]	Lytro Illum	University of Rome	2016
Disney LFs ^[51]	A DSLR camera on a motorized linear stage	Disney	2013
UCB LF ^[79]	Synthetic Lytro Illum	University of California,Berkeley(UCB)	2015
DDFT ^[99]	Lytro Illum	Technische Universität München(TUM)	2018
LF and Stereo ^[100]	Raytrix R5 Stereo camera		2018
INRIA LF dataset ^[101]	Lytro Lytro Illum Raytrix R8 Synthetic	Institut national de recherche en informatique et en automatique (INRIA)	2015, 2018, 2018, 2019
MLRD ^[102]	Lytro Illum Raytrix R29	Mid Sweden University	2018
CPCD ^[103]	Custom-built	University of Bedfordshire	2016

表 5 与光场三维重建相关的开源算法小结

用途	方法	发布机构	时间
光场相机标定	文献[61]	University of Sydney	2013
	文献[62]	Korea Advanced Institute of Science and Technology(KAIST)	2017
	文献[63]	西北工业大学 上海科技大学	2019
	文献[70]	University College London	2017
	文献[67]	海军航空大学	2020
光场影像三维深度恢复	文献[80]	University of Delaware; Microsoft Reserach	2014
	文献[75]	KAIST	2015
	文献[79]	UCB University of California,San Diego(UCSD)	2015
	文献[31]	Bina Nusantara University Inha University Seoul National University	2018
	文献[86]	北京航空航天大学 University of Wisconsin-Milwaukee	2016
	文献[89]	Yonsei University KAIST	2018
	文献[90]	台湾大学 MediaTek	2020

6 总结

光场相机作为一种新型相机,能够通过单次拍摄同时采集空间和方向信息,为三维重建提供了新的数据源.目前,光场相机的三维重建研究处于探索阶段,一些基本问题还尚未被解决.为挖掘光场相机的应用潜力,本文对国内外围绕光场相机的三维重建展开的诸多研究工作进行了梳理和总结,并指出了目前存在的主要问题和未来可能的研究方向.

(1) 当前的光场相机多为国外生产,迫切需要研制自主可控的国产光场相机.由于空间分辨率和方向分辨率的相互矛盾,相机参数的制定过程中应根据应用进行专业化设计,在提高方向分辨率的同时,保证空间分辨率不会明显下降,并提高拍照速度.

(2) 新型光场相机的出现必然会给相机标定引入新的问题,而准确有效地完成相机标定是进行下一步工作的基础.为得到准确通用的标定模型,构建投影模型时需考虑不同内部设计的光场相机并兼顾不同镜头的相机畸变.同时,在标定完成后,需构建大量不同条件的光场影像数据集,对标定算法进行全面而深入的评估.

(3) 光场影像固有地包含多视角信息,可用以评估影像的深度.虽然部分深度估计算法在大多数区域都能取得较好的效果,但在遮挡的区域表现较差.为了去除遮挡的影响,其中一个极端是构建同时考虑所有视角遮挡和深度的完全遮挡评估模型,但该方法效率较低.另一个极端是,利用某种策略预先估计被遮挡视角,并利用剩下未被遮挡的视角来评估深度.目前大部分的方法都是基于第二种策略,该类方法难以保证剔除的视角都是被遮挡的视角,被错误剔除的视角信息就损失了,在后续的优化过程中,这些被错误剔除的信息也没有被应用到.因此,在评估过程以及优化过程,从整体和局部同时考虑遮挡是未来的主要研究方向.此外,基于学习的三维深度恢复方法表现出巨大的潜力.

总体来说,光场相机给三维重建提供新的思路的同时,也引入了诸多问题.如何准确地恢复相机内部的光场,实现场景的精细化三维重建在未来的一段时间依然是研究的热点问题,还需要更进一步地深入研究.

参考文献

- [1] 王之卓. 摄影测量原理[M]. 北京: 测绘出版社, 1990.
WANG Zhi-zuo. Principle of Photogrammetry[M]. Beijing: Publishing House of Surveying and Mapping, 1990. (in Chinese)
- [2] 张祖勋, 张剑清, 张力. 数字摄影测量发展的机遇与挑战[J]. 武汉测绘科技大学学报, 2000, 25(1): 7-11.
ZHANG Zu-xun, ZHANG Jian-qing, ZHANG Li. Opportunities and challenges for development of digital photogrammetry[J]. Journal of Wuhan Technical University of Surveying and Mapping(Wtusm), 2000, 25(1): 7-11. (in Chinese)
- [3] NEX F, REMONDINO F. UAV for 3D mapping applications: A review[J]. Applied Geomatics, 2014, 6(1): 1-15.
- [4] BRUNO F, BRUNO S, DE SENSI G, et al. From 3D reconstruction to virtual reality: A complete methodology for digital archaeological exhibition[J]. Journal of Cultural Heritage, 2010, 11(1): 42-49.
- [5] YANG M D, CHAO C F, HUANG K S, et al. Image-based 3D scene reconstruction and exploration in augmented reality[J]. Automation in Construction, 2013, 33: 48-60.
- [6] NUCHTER A, SURMANN H, HERTZBERG J. Automatic model refinement for 3D reconstruction with mobile robots[C]//Fourth International Conference on 3-D Digital Imaging and Modeling, 2003.3DIM 2003. Banff, Canada: IEEE, 2003: 394-401.
- [7] SONG L M, LI X Y, YANG Y G, et al. Structured-light based 3D reconstruction system for cultural relic packaging[J]. Sensors, 2018, 18(9): 2981.
- [8] JIANG D L, HU Y X, YAN S C, et al. Efficient 3D reconstruction for face recognition[J]. Pattern Recognition, 2005, 38(6): 787-798.
- [9] KHAN U, YASIN A U, ABID M, et al. A methodological review of 3D reconstruction techniques in tomographic imaging[J]. Journal of Medical Systems, 2018, 42(10): 190.
- [10] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces[C]//2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality. Nara, Japan: IEEE, 2007: 225-234.
- [11] DAVISON A J, REID I D, MOLTON N D, et al. MonoSLAM: Real-time single camera SLAM[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(6): 1052-1067.
- [12] MUR-ARTAL R, MONTIEL J M M, TARDÓS J D. ORB-SLAM: A versatile and accurate monocular SLAM system[J]. IEEE Transactions on Robotics, 2015, 31(5): 1147-1163.
- [13] MUR-ARTAL R, TARDÓS J D. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras[J]. IEEE Transactions on Robotics, 2017, 33(5): 1255-1262.
- [14] CAMPOS C, ELVIRA R, RODRÍGUEZ J J G, et al. ORB-SLAM3: An accurate open-source library for visual, visual-inertial, and multimap SLAM[J]. IEEE Transactions on Robotics, 2021, 37(6): 1874-1890.
- [15] HESS W, KOHLER D, RAPP H, et al. Real-time loop closure in 2D LIDAR SLAM[C]//2016 IEEE International

- al Conference on Robotics and Automation. Stockholm, Sweden: IEEE, 2016: 1271-1278.
- [16] ZHANG J, SINGH S. LOAM: Lidar odometry and mapping in real-time[C]//Robotics: Science and Systems. Berkeley, California, USA: University of California, Berkeley, 2014, 2(9): 1-9.
- [17] YE H Y, CHEN Y Y, LIU M. Tightly coupled 3D lidar inertial odometry and mapping[C]//2019 International Conference on Robotics and Automation(ICRA). Montreal, Canada: IEEE, 2019: 3144-3150.
- [18] KOIDE K, MIURA J, MENEGATTI E. A portable three-dimensional LIDAR-based system for long-term and wide-area people behavior measurement[J]. International Journal of Advanced Robotic Systems, 2019, 16(2): 1729881419841532.
- [19] XU W, ZHANG F. FAST-LIO: A fast, robust LiDAR-inertial odometry package by tightly-coupled iterated Kalman filter[J]. IEEE Robotics and Automation Letters, 2021, 6(2): 3317-3324.
- [20] CHEN X, MILIOTO A, PALAZZOLO E, et al. SuMa++: Efficient LiDAR-based semantic SLAM[C]//2019 IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS). Macau, China: IEEE, 2019: 4530-4537.
- [21] KERL C, STURM J, CREMERS D. Dense visual SLAM for RGB-D cameras[C]//2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. Tokyo, Japan: IEEE, 2013: 2100-2106.
- [22] NEWCOMBE R A, IZADI S, HILLIGES O, et al. Kinect-Fusion: Real-time dense surface mapping and tracking [C]//2011 10th IEEE International Symposium on Mixed and Augmented Reality. Basel, Switzerland: IEEE, 2011: 127-136.
- [23] WHELAN T, LEUTENEGGER S, SALAS MORENO R, et al. ElasticFusion: Dense SLAM without a pose graph [C]//Robotics: Science and Systems. Rome, Italy: Sapienza University of Rome, 2015: 1-9.
- [24] DAI A, NIEBNER M, ZOLLHÖFER M, et al. BundleFusion: Real-time globally consistent 3D reconstruction using on-the-fly surface reintegration[J]. ACM Transactions on Graphics, 2017, 36(4): 76a.
- [25] SU Z, XU L, ZHENG Z R, ET AL. RobustFusion: Human Volumetric Capture with Data-driven Visual Cues Using A RGBD Camera[M]//Computer Vision-ECCV 2020. Switzerland: Springer, Cham, 2020: 246-264.
- [26] NG R. Light field photography with a hand-held plenoptic camera[D]. Palo Alto: Stanford University, 2005.
- [27] NG R. Fourier slice photography[J]. ACM Transactions on Graphics, 2005, 24(3): 735-744.
- [28] GMBH RAYTRIX. Raytrix3d light field camera technology[EB/OL]. [2022]. <http://www.raytrix.de/>.
- [29] VENKATARAMAN K, LELESCU D, DUPARRÉ J, et al. PiCam: An ultra-thin high performance monolithic camera array[J]. ACM Transactions on Graphics, 2013, 32(6): 1-13.
- [30] 方璐, 戴琼海. 计算光场成像[J]. 光学学报, 2020, 40(1): 9-30.
FANG Lu, DAI Qing-hai. Computational light field imaging[J]. Acta Optica Sinica, 2020, 40(1): 9-30. (in Chinese)
- [31] WILLIEM, PARK I K, LEE K M. Robust light field depth estimation using occlusion-noise aware data costs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(10): 2484-2497.
- [32] ZELLER N, QUINT F, STILLA U. Scale-awareness of light field camera based visual odometry[C]//Proceedings of the European Conference on Computer Vision(ECCV). Munich, Germany: Springer, 2018: 715-730.
- [33] MITRA K, VEERARAGHAVAN A. Light field denoising, light field superresolution and stereo camera based refocussing using a GMM light field patch prior[C]//2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Providence, RI, USA: IEEE, 2012: 22-28.
- [34] 孙福盛, 韩燮. 光场相机精确色彩矢量约束下的超分辨率算法[J]. 光学学报, 2019, 39(3): 59-68.
SUN Fu-sheng, HAN Xie. Super-resolution algorithm based on precise color vector constraint of light field camera[J]. Acta Optica Sinica, 2019, 39(3): 59-68. (in Chinese)
- [35] 刘亚美, 张骏, 张旭东, 等. 光场显著性检测研究综述 [J]. 中国图象图形学报, 2020, 25(12): 2465-2483.
LIU Ya-mei, ZHANG Jun, ZHANG Xu-dong, et al. Review of saliency detection on light fields[J]. Journal of Image and Graphics, 2020, 25(12): 2465-2483. (in Chinese)
- [36] YÜCER K, SORKINE-HORNUNG A, WANG O, et al. Efficient 3D object segmentation from densely sampled light fields with applications to 3D reconstruction[J]. ACM Transactions on Graphics, 2016, 35(3): 1-15.
- [37] PEI Z, ZHANG Y N, YANG T, et al. A novel multi-object detection method in complex scene using synthetic aperture imaging[J]. Pattern Recognition, 2012, 45(4): 1637-1658.
- [38] GUO X Q, YU Z, KANG S B, et al. Enhancing light fields through ray-space stitching[J]. IEEE Transactions on Visualization and Computer Graphics, 2016, 22(7): 1852-1861.

- [39] 刘玉轩. 光场相机三维重建方法研究[D]. 武汉:武汉大学, 2020.
- LIU Yu-xuan. 3D Reconstruction Based on Light Field Cameras[D]. Wuhan: Wuhan University, 2020. (in Chinese)
- [40] WU G C, MASIA B, JARABO A, et al. Light field image processing: An overview[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2017, 11(7): 926-954.
- [41] GERSHUN A. The light field[J]. *Journal of Mathematics and Physics*, 1939, 18(1/2/3/4): 51-151.
- [42] ADELSON E H, BERGEN J R. The plenoptic function and the elements of early vision[J]. *Computational Models of Visual Processing*, 1991, 3-20.
- [43] LEVOY M, HANRAHAN P. Light field rendering[C]// *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques-SIGGRAPH'96*. New York: ACM Press, 1996: 31-42.
- [44] JOHANNSEN O, HONAUER K, GOLDLUECKE B, et al. A taxonomy and evaluation of dense light field depth estimation algorithms[C]// *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Honolulu, HI, USA: IEEE, 2017: 1795-1812.
- [45] YANG J C, EVERETT M, BUEHLER C, et al. A real-time distributed light field camera[J]. *Rendering Techniques*, 2002, 2002: 77-86.
- [46] WILBURN B S, SMULSKI M, LEE H H K, et al. Light field video camera[C]// *Proceedings of Media Processors 2002*. San Jose: SPIE, 2001, 4674: 29-36.
- [47] WILBURN B, JOSHI N, VAISH V, et al. High performance imaging using large camera arrays[J]. *ACM Transactions on Graphics*, 2005, 24(3): 765-776.
- [48] ZHANG C, CHEN T. A self-reconfigurable camera array [C]// *SIGGRAPH'04: ACM SIGGRAPH 2004 Sketches*. New York: ACM Press, 2004: 151.
- [49] LIN X, WU J M, ZHENG G A, et al. Camera array based light field microscopy[J]. *Biomedical Optics Express*, 2015, 6(9): 3179-3189.
- [50] VAISH V, ADAMS A. Stanford(New) light field archive [EB/OL].[2022]. <http://lightfield.stanford.edu/lfs.html>.
- [51] KIM C, ZIMMER H, PRITCH Y, et al. Scene reconstruction from high spatio-angular resolution light fields[J]. *ACM Transactions on Graphics*, 2013, 32(4): 1-12.
- [52] LIANG C K, LIN T H, WONG B Y, et al. Programmable aperture photography[J]. *ACM Transactions on Graphics*, 2008, 27(3): 1-10.
- [53] VEERARAGHAVAN A, RASKAR R, AGRAWAL A, et al. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing[J]. *ACM Transactions on Graphics*, 2007, 26(3): 69.
- [54] MARWAH K, WETZSTEIN G, BANDO Y, et al. Compressive light field photography using overcomplete dictionaries and optimized projections[J]. *ACM Transactions on Graphics*, 2013, 32(4): 46(1-12).
- [55] ADELSON E H, WANG J Y A. Single lens stereo with a plenoptic camera[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, 14(2): 99-106.
- [56] DUAN H X, MEI L, WANG J, et al. A new imaging model of Lytro light field camera and its calibration[J]. *Neurocomputing*, 2019, 328: 189-194.
- [57] LUMSDAINE A, GEORGIEV T. The focused plenoptic camera[C]// *2009 IEEE International Conference on Computational Photography*. San Francisco: IEEE, 2009: 1-8.
- [58] LEVOY M, NG R, ADAMS A, et al. Light field microscopy[J]. *ACM Transactions on Graphics*, 2006, 25(3): 924-934.
- [59] SHEHZAD K, XU Y. Graphene light-field camera[J]. *Nature Photonics*, 2020, 14(3): 134-136.
- [60] PLESS R. Using many cameras as one[C]// *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. WI, USA: IEEE, 2003: II-587.
- [61] DANSEREAU D G, PIZARRO O, WILLIAMS S B. Decoding, calibration and rectification for lenselet-based plenoptic cameras[C]// *2013 IEEE Conference on Computer Vision and Pattern Recognition*. Portland, Oregon, USA: IEEE, 2013: 1027-1034.
- [62] BOK Y, JEON H G, KWEON I S. Geometric calibration of micro-lens-based light field cameras using line features [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(2): 287-300.
- [63] ZHANG Q, ZHANG C P, LING J B, et al. A generic multi-projection-center model and calibration method for light field cameras[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(11): 2539-2552.
- [64] JI Y, WU J. Calibration method of light-field camera for photogrammetry application[J]. *Measurement*, 2019, 148: 106943.
- [65] PERTUZ S, PULIDO-HERRERA E, KAMARAINEN J K. Focus model for metric depth estimation in standard plenoptic cameras[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 144: 38-47.
- [66] ZHOU P, YANG Z, CAI W J, et al. Light field calibration and 3D shape measurement based on epipolar-space [J]. *Optics Express*, 2019, 27(7): 10171-10184.
- [67] LIU Q S, XIE X F, ZHANG X Z, et al. Stepwise calibration of plenoptic cameras based on corner features of raw images[J]. *Applied Optics*, 2020, 59(14): 4209-4219.

- [68] 刘青松, 谢晓方, 张烜喆, 等. 用于聚焦型光场相机标定的棋盘角点检测算法[J]. 光学学报, 2020, 40(14): 153-160.
- LIU Qing-song, XIE Xiao-fang, ZHANG Xuan-zhe, et al. Checkboard corner detection algorithm for calibration of focused plenoptic camera[J]. Acta Optica Sinica, 2020, 40(14): 153-160. (in Chinese)
- [69] LIU Y X, MO F, ALEKSANDROV M, et al. Accurate calibration of standard plenoptic cameras using corner features from raw images[J]. Optics Express, 2021, 29(1): 158-169.
- [70] NOUSIAS S, CHADEBECQ F, PICHAT J, et al. Corner-based geometric calibration of multi-focus plenoptic cameras[C]//2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 957-965.
- [71] YU J Y, MCMILLAN L, GORTLER S. Seam light field rendering[C]//10th Pacific Conference on Computer Graphics and Applications. Beijing, China: IEEE, 2002: 137-144.
- [72] BOLLES R C, BAKER H H, MARIMONT D H. Epipolar-plane image analysis: An approach to determining structure from motion[J]. International Journal of Computer Vision, 1987, 1(1): 7-55.
- [73] HEBER S, RANFTL R, POCK T. Variational shape from light field[C]//Energy Minimization Methods in Computer Vision and Pattern Recognition. Berlin: Springer, 2013: 66-79.
- [74] YU Z, GUO X Q, LING H B, et al. Line assisted light field triangulation and stereo matching[C]//2013 IEEE International Conference on Computer Vision. Sydney: IEEE, 2013: 2792-2799.
- [75] JEON H G, PARK J, CHOE G, et al. Accurate depth map estimation from a lenslet light field camera[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 1547-1555.
- [76] TAO M W, HADAP S, MALIK J, et al. Depth from combining defocus and correspondence using light-field cameras[C]//2013 IEEE International Conference on Computer Vision. Sydney: IEEE, 2013: 673-680.
- [77] TAO M W, SRINIVASAN P P, MALIK J, et al. Depth from shading, defocus, and correspondence using light-field angular coherence[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 1940-1948.
- [78] WANG T C, EFROS A A, RAMAMOORTHY R. Depth estimation with occlusion modeling using light-field cameras[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(11): 2170-2181.
- [79] WANG T C, EFROS A A, RAMAMOORTHY R. Occlusion-aware depth estimation using light-field cameras[C]//2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 3487-3495.
- [80] CHEN C, LIN H T, YU Z, et al. Light field stereo matching using bilateral statistics of surface cameras[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, Ohio, USA: IEEE, 2014: 1518-1525.
- [81] WILLIEM W, PARK I K. Robust light field depth estimation for noisy scene with occlusion[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 4396-4404.
- [82] WANNER S, GOLDLUECKE B. Variational light field analysis for disparity estimation and super-resolution[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(3): 606-619.
- [83] LI J Q, LU M L, LI Z N. Continuous depth map reconstruction from light fields[J]. IEEE Transactions on Image Processing, 2015, 24(11): 3257-3265.
- [84] ZHANG Y B, LV H J, LIU Y B, et al. Light-field depth estimation via epipolar plane image analysis and locally linear embedding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27(4): 739-747.
- [85] SUZUKI T, TAKAHASHI K, FUJII T. Sheared EPI analysis for disparity estimation from light fields[J]. IEEE Transactions on Information and Systems, 2017, E100.D(9): 1984-1993.
- [86] ZHANG S, SHENG H, LI C, et al. Robust depth estimation for light field via spinning parallelogram operator[J]. Computer Vision and Image Understanding, 2016, 145: 148-159.
- [87] SHENG H, ZHAO P, ZHANG S, et al. Occlusion-aware depth estimation for light field using multi-orientation EPIs[J]. Pattern Recognition, 2018, 74: 587-599.
- [88] SCHILLING H, DIEBOLD M, ROTHER C, et al. Trust your model: Light field depth estimation with inline occlusion handling[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 4530-4538.
- [89] SHIN C, JEON H G, YOON Y, et al. EPINET: A fully-convolutional neural network using epipolar geometry for depth from light field images[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 4748-4757.
- [90] TSAI Y J, LIU Y L, OUHYOUNG M, et al. Attention-based view selection networks for light-field disparity estimation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York: AAAI Press, 2020, 34(7):

- 12095-12103.
- [91] ZHOU W H, LIANG L K, ZHANG H, et al. Scale and orientation aware EPI-patch learning for light field depth estimation[C]//2018 24th International Conference on Pattern Recognition(ICPR). Beijing: IEEE, 2018: 2362-2367.
- [92] WANG X C, TAO C N, WU R M, et al. Light-field-depth-estimation network based on epipolar geometry and image segmentation[J]. Journal of the Optical Society of America. A, 2020, 37(7): 1236-1243.
- [93] WANNER S, MEISTER S, GOLDLUECKE B. Datasets and benchmarks for densely sampled 4D light fields[C]// Proceedings of 18th VMV. Lugano, Switzerland: Università della Svizzera italiana(USI), 2013, 13: 225-226.
- [94] RERABEK M, EBRAHIMI T. New light field image dataset[C]//Proceedings of the 8th International Conference on Quality of Multimedia Experience(QoMEX). Lisbon, Portugal: IEEE, 2016: 1-2.
- [95] DANSEREAU D G, GIROD B, WETZSTEIN G. LiFF: Light field features in scale and depth[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach: IEEE, 2019: 8034-8043.
- [96] WETZSTEIN G. The MIT Synthetic Light Field Archive [EB/OL]. [2022-02-22]. <https://web.media.mit.edu/~gordonw/SyntheticLightFields/index.php>.
- [97] HONAUER K, JOHANNSEN O, KONDERMANN D, et al. A dataset and evaluation methodology for depth estimation on 4D light fields[C]//Computer Vision-ACCV 2016. Taipei: Springer, Cham, 2017: 19-34.
- [98] PAUDYAL P, OLSSON R, SJÖSTRÖM M, et al. SMART: A light field image quality dataset[C]//MM-Sys' 16: Proceedings of the 7th International Conference on Multimedia Systems. Klagenfurt, Austria: ACM Press, 2016: 1-6.
- [99] HAZIRBAS C, SOYER S G, STAAB M C, et al. Deep depth from focus[C]//Proceedings of 2018 Asian Conference on Computer Vision. Perth, Australia: Springer, Cham, 2018: 525-541.
- [100] ZELLER N, QUINT F, STILLA U. A synchronized stereo and plenoptic visual odometry dataset[DB/OL]. [2018]. <http://arXiv:1807.09372>.
- [101] INRIA. Light field software & datasets[DS/OL]. [2022-02-22]. <http://clim.inria.fr/DataSoftware.html>.
- [102] AHMAD W, PALMIERI L, KOCH R, et al. Matching light field datasets from plenoptic cameras 1.0 and 2.0 [C]//2018-3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video(3DTV-CON). Helsinki, Finland: IEEE, 2018: 1-4.
- [103] HAHNE C, AGGOUN A, VELISAVLJEVIC V, et al.

Refocusing distance of a standard plenoptic camera[J]. Optics Express, 2016, 24(19): 21521-21540.

作者简介



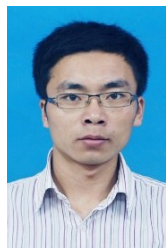
刘玉轩 男,1990年生于河南省南阳市. 现为中国测绘科学研究院摄影测量与遥感研究所助理研究员. 主要研究方向为摄影测量与计算机视觉、光场相机三维重建等.
E-mail: yxliu@casm.ac.cn



张力 男,1970年生于山东省潍坊市. 现为中国测绘科学研究院摄影测量与遥感研究所研究员. 主要研究方向为数字摄影测量、卫星数据处理、点云语义化建模等.
E-mail: zhangl@casm.ac.cn



艾海滨 男,1976年生于江西省九江市. 现为中国测绘科学研究院摄影测量与遥感研究所研究员. 主要研究方向为数字摄影测量、航空影像三维建模等.
E-mail: Aihb@casm.ac.cn



许彪(通讯作者) 男,1986年生于吉林省松原市. 现为中国测绘科学研究院摄影测量与遥感研究所副研究员. 主要研究方向为航空影像三维重建、计算机视觉等.
E-mail: biaoxx@casm.ac.cn



孙钰珊 女,1982年生于黑龙江省哈尔滨市. 现为中国测绘科学研究院摄影测量与遥感研究所助理研究员. 主要研究方向为卫星影像处理、无人机影像三维重建等.
E-mail: sunys@casm.ac.cn



樊仲霖 男,1997年生于山东省菏泽市. 现为中国测绘科学研究院摄影测量与遥感研究所硕士研究生. 主要研究方向为星载多源数据联合处理.
E-mail: fz110466@163.com