

· 特邀综述 ·

计算光场成像

方璐, 戴琼海*

清华大学脑与认知科学研究院, 北京 100084

摘要 光场为三维世界中光线集合的完备表示。通过记录更高维度的光线数据, 光场能够准确感知周围复杂多变的动态环境, 支撑智能系统对环境的理解与决策。计算光场成像技术围绕光场及全光函数表示, 旨在结合计算、数字传感器、光学系统和智能光照等技术, 以及硬件设计、软件计算能力, 突破经典成像模型和数字相机的局限性, 建立光在空域、视角、光谱和时域等多个维度的关系, 实现耦合感知、解耦重建与智能处理, 具备面向大范围动态场景的多维多尺度成像能力。光场成像技术正逐渐被应用于生命科学、工业探测、国家安全、无人系统和虚拟现实/增强现实等领域, 具有重要的学术研究价值和广阔的产业应用前景。然而, 伴随着高维数据的离散化采样, 光场成像面临空间分辨率与视角分辨率的维度权衡挑战, 如何对稀疏化的采样数据进行光场重建成为计算光场成像及其应用的基础难题。与此同时, 受制于光场信号的高维数据感知量, 光场处理面临有效数据感知与计算高效性的矛盾。如何用光场这一高维信息采集手段, 取代传统二维成像视觉感知方法, 并结合智能信息处理技术实现智能化高效感知, 是实现光场成像技术产业化应用的巨大挑战。对过去 20 年来计算光场成像装置与算法的研究进行概述和讨论, 内容涵盖光场表示和理论、光场信号采集、空间与视角维度重建等。

关键词 成像系统; 计算摄像; 光场成像; 十亿像素; 阵列相机

中图分类号 O438

文献标志码 A

doi: 10.3788/AOS202040.0111001

Computational Light Field Imaging

Fang Lu, Dai Qionghai*

Institute for Brain and Cognitive Sciences, Tsinghua University, Beijing 100084, China

Abstract High performance imaging of large-scale dynamic scenes is substantial to vision intelligence. The light field is a 3D plenoptic function that describes the amount of light flow in every direction through every point in space. By recording the high dimensional light signal, the light field can accurately perceive the complex dynamic environment, supporting the understanding and decision-making of the intelligent system. Computational light field imaging technique, based on the light field and the representation of plenoptic function, aims to combine computation, digital sensors, optical system, and intelligent lighting, thereby combining the hardware design and software computing power. This technique breaks through the limits of classical imaging model and digital camera, establishes the relationship among light in spatial, angular, spectral, and temporal dimensions, realizes coupling perception, decoupling reconstruction, and intelligent processing, and leads to the multi-dimensional and multi-scale imaging ability for large-scale dynamic scenes. Light field imaging technique plays vital role in various fields, including life science, industrial inspection, national security, unmanned system, VR/AR, etc., attracting broad interests from both academia and industry. With the discrete sampling of high dimensional data, light field imaging faces the challenge of dimension trade-off between spatial resolution and angular resolution. How to reconstruct light field for sparse sampled data becomes a fundamental problem in computational light field imaging and its applications. Meanwhile, limited by high dimensional data perception of light field signals, light field process faces the contradiction between effective data perception and computational efficiency. How to replace the traditional two-dimensional imaging visual perception method with light field which is a high-dimensional information acquisition means and how to combine intelligent information processing technique to realize intelligent efficient perception, are huge challenges for industrial applications of the light field imaging technique. In this paper, we conduct a thorough literature review of devices and algorithms of computational light field imaging, including the representation and theory of light field, light field signal sampling, and light field reconstruction with super-resolution in spatial and angular domain.

收稿日期: 2019-11-09; 修回日期: 2019-12-17; 录用日期: 2019-12-26

* E-mail: daiqionghai@tsinghua.edu.cn

Key words imaging systems; computational photography; light field imaging; gigapixel; array camera

OCIS codes 110.1758; 110.6895

1 引 言

光是人类对物理世界进行观测与感知最重要的载体之一,人类通过人眼接收场景中物体发出的光线(主动或被动发光)进行感知。成像感知系统是人眼的延伸,捕获、记录、分析场景的光信息。然而,现有的成像感知系统大多仅支持二维成像,迫使我们只能通过二维窗口去观察三维世界,从而丢失了三维世界的丰富信息。三维世界中光线集合的完备表示被称为光场。与传统二维成像将光线直接投影在平面感光元件上不同,光场成像通过捕捉光线在自由空间中的分布来获取场景的三维信息。因此,光场通过记录更高维度的光线数据,能够获取比传统二维成像以及以双目立体视觉为代表的传统三维成像更高精度的三维信息,从而准确感知周围复杂多变的动态环境,支撑智能系统对环境的理解与智能决策。

描述光线分布的早期模型由 Gershun^[1]于 1939 年定义,在这篇经典的文章中,“Light Field”作为专业术语被首次提出。然而,当时“Light Field”描述的概念主要是指空间中的光辐射可以表示为关于空间位置的三维向量。虽然“Light Field”在当时描述的仅仅是光辐射,但是从命名来看,该术语可以描述更广阔的光学领域。在半个世纪之后的 1981 年, Moon 等^[2]提出了“Photic Field”的概念,相比于前人的“Light Field”,此为当前学术领域研究的光场,该概念于 1991 年由 Adelson 等^[3]进一步完善。该光线分布模型被称为全光函数,其将场景中的光描述为空间位置 (x, y, z) 、角度 (θ, φ) 、波长和时间的七维函数。然而,这种七维全光函数采集的数据量巨大,当时的计算机技术水平难以实现如此庞大数据的实时传输与处理,因此,Levoy 等^[4]于 1996 年在假设光线沿光传播方向辐射度不变的基础上,提出了四维光场的概念,即使用光线的位置和方向来描述光场,极大地简化了光场数据的获取。值得注意的是,在这篇经典的论文中,用来描述空间光线分布的概念再次由前人的“Photic Field”改为“Light Field”。

简化后的四维光场并不能完备地表示三维空间中的所有光线,但是可以完备地描述人眼接收的光线。计算光场成像技术基于该四维光场表示,旨在

建立光在空域、视角、光谱和时域等多个维度的关系,实现耦合感知、解耦重建与智能处理,用于面向大范围动态场景的多维多尺度成像^[4]。光场成像技术正逐渐应用于生命科学、工业探测、国家安全、无人系统和虚拟现实/增强现实等领域,具有重要的学术研究价值和产业应用前景。以目前广泛应用于智能系统中的成像单元为例,其大多基于单目或双目像传感器,受采样维度和空间带宽积的物理限制,难以超越人眼所具备的多维多尺度成像能力,在时间、空间、角度、动态范围等多个维度上的成像性能严重受限,如单目摄像头内窥镜无法获取肿瘤的深度信息,需要多角度观察,极大地降低了手术效率。为探索智能成像的新方向,亟需开展基于像传感器阵列计算光场成像的研究,实现面向大范围动态场景的多维多尺度感知手段。鉴于光场巨大的使用潜力以及待攻克的技术难题,光场技术受到了国内外研究团队的广泛关注。在近 30 年来,光场理论逐步得到了完善,形成了统一的理论体系,尤其是在光场的采集和显示方面。

现代光场^[4]主要关注光线作为位置和角度的分布函数。受益于空间-视角维度假设以及软硬件技术的进步,如今,光场成像可由便携化的商用设备实现^[5],甚至可集成到手机中^[6]。此外,它们亦可被整合到显微镜^[7]中,实现快速扫描下生物样本的三维成像。光场相机的广泛使用使一些新的应用成为可能,包括最初的基于真实感图像的光场渲染以及现在利用光场编码信息的计算机视觉应用,例如 3D 重建、分割、显著性检测、物体检测与识别、跟踪和视频增稳等。

尽管光场研究在过去 30 年取得了重要进展,但在采集、重建、显示等方面仍面临诸多挑战。从采集的角度来看,光场成像亟需具备捕获高维数据的能力,即应在传感器上捕获每个位置上的定向光分布,而不是简单地在传统摄影中记录 2D 投影。然而,为获取高维数据,通常会在各维度之间进行分辨率的权衡。目前常用的光场成像方式主要有两种^[8-10]:一是利用相机阵列采集来自不同方向(分布)的光线;二是将不同方向的光线信息与不同位置的光线信息混合编码至平面成像的芯片中。前者由于相机物理尺寸的限制,相机间始终存在较大的基线距离,因而无法采集密集角度(视点)的信息;而

后者因受成像芯片的像素分辨率限制,无法采集密集空间维度(即传统意义上的图像分辨率)的信息。在实际应用中,尤其是利用相机阵列采集光场数据时,光场的视点采样往往是稀疏的、欠采样的,因此如何利用稀疏化的采样数据进行光场重建是光场成像及其应用的基础难题。此外,光场的高维度极大地增加了对其进行处理和分析的复杂度,在算法设计上施加了更具挑战性的条件。例如,传统 2D 图像中的分割旨在将单个图像内的前景和背景分开,而在光场中执行分割时,不仅要多个视点的图像进行处理,还需要保持光场结构的稳定性。如何用光场这一高维信息采集手段,取代传统二维成像这一视觉感知方法,并结合智能信息处理技术实现智能化感知功能,是实现光场成像技术产业应用面临的巨大挑战。

本文旨在对计算光场成像和处理领域的相关研究进行概述,同时揭示其中的挑战并汇总解决方案。首先介绍了光场的理论描述(第 2 节);随后深入讨论了光场成像的三个细分领域,低层级上基于硬件的采集,重点关注有效捕获光场所需的权衡(第 3 节),高层上的处理则涵盖空间(第 4 节)和视角(第 5 节)维度高分辨率成像的研究进展;最后针对

计算光场成像进行了总结与展望(第 6 节)。

2 光场理论

本节简要介绍光场模型,包括光场的多维函数表征以及光场的可视化和渲染。

2.1 光场表示

全光函数^[2]是一个七维函数 $L(x, y, z, \theta, \varphi, \lambda, t)$,用以描述三维空间中每个点在每个方向上传播的光线组。为获得此函数,在每个角度 (θ, φ) 、每个波长 λ 和每个时间 t 上测量每个位置 (x, y, z) 处的光线辐射度。然而,受制于高维数据记录和处理等难题,必须对实际使用的光场模型进行简化。首先,假设被测函数具有单色性和时不变性(动态光场将随着离散时间序列 t 记录在光场视频的不同帧中),即认为光线不随时间变化,且其波长 λ 被独立记录在不同的颜色通道中。在这样的假设下,我们可以从全光函数中去除波长 λ 和时间 t 维度,从而将模型从七维降到五维。其次,Levoy 等^[4]对全光函数进行了第二次简化,他们意识到五维表示仍然包含一些冗余数据,若假设光线是在自由空间中传播,光线辐射沿传播直线保持不变,便能进一步将五维表示降维到四维表示,如图 1 所示。

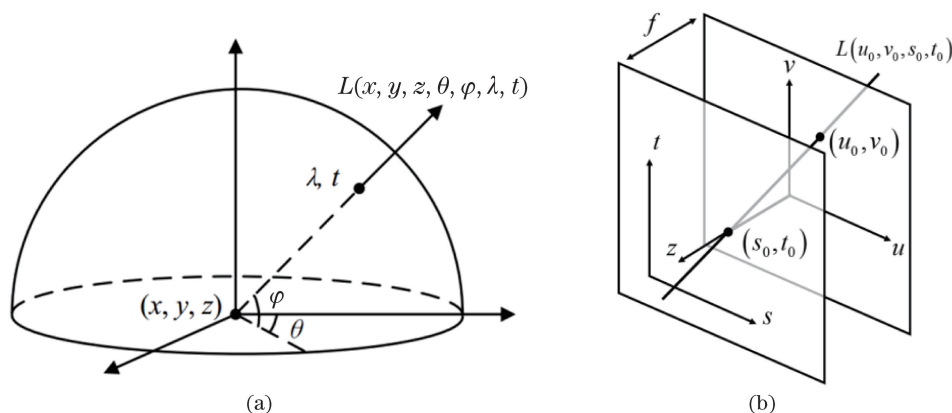


图 1 (a)七维全光函数和(b)四维简化光场示意图

Fig. 1 Schematics of (a) the seven-dimensional plenoptic function and (b) the four-dimensional simplified light field

在对四维光场进行参数化时需考虑三个关键问题^[3]:计算效率、对光线集的控制以及光场空间的均匀采样。基于这些问题,四维光场表示的最常见解决方案是设置两个平行平面,然后用光线与两平面相交点的坐标来进行参数化。对于第一个平面,坐标系由 (u, v) 表示,对于第二个平面,坐标系由 (s, t) 表示。系统中定义的定向光线首先在 (u, v) 坐标处与 uv 平面相交,然后在坐标 (s, t) 处与 st 平面相交,这样的光线记作 $L(u, v, s, t)$ 。这样,描述

光场的全光函数就从七维降到了四维,并由四个坐标 (u, v, s, t) 进行参数化。

在双平面光场模型 $L(u, v, s, t)$ 中,我们可将 st 平面视为一组摄像机,其焦平面位于 uv 平面上。双平面表示法可以通过两种不同的角度来理解。其一,每个摄像机收集离开 st 平面并到达 uv 平面焦点的光线(某个视点接收的光线集合)。因此,四维光场可表示为二维图像阵列,例如图 2 中所示的图像集合。由相机记录的每个图像被称为子孔径图

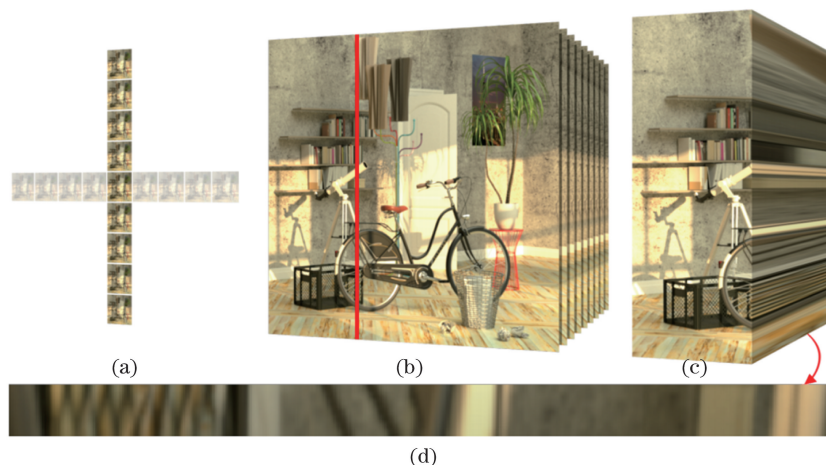


图 2 四维光场数据的表征。(a)子孔径图像 $I_{s^*, t^*}(u, v)$, 可通过固定 st 坐标(s^* 和 t^*)对光场采样来获取;(b)通过固定 uv 坐标(u^* 和 v^*)采样获取的光场子视图 $I_{u^*, v^*}(s, t)$;(c) 通过固定 u 和 s 获得的极线平面图像 $E_{u^*, s^*}(v, t)$;
(d)通过固定 v 和 t 获得的极线平面图像 $E_{v^*, t^*}(u, s)$

Fig. 2 Characterization of the four-dimensional light field. (a) Sub-aperture image $I_{s^*, t^*}(u, v)$, obtained by sampling in the light field with fixed st coordinates (s^* and t^*); (b) sub-image of the light field $I_{u^*, v^*}(s, t)$ by sampling with fixed uv coordinates (u^* and v^*); (c) EPI $E_{u^*, s^*}(v, t)$ by fixing u and s ; (d) EPI $E_{v^*, t^*}(u, s)$ by fixing v and t

像,也被称为针孔视图。其二, uv 平面上的点与一簇离开该点并到达 st 平面的光线相绑定。因为 st 平面中的采样数取决于视点的数量,而 uv 平面中的采样数取决于相机分辨率, s 和 t 维度称为角度维度, u 和 v 维度称为空间维度。

具体地,对于四维光场 $L(u, v, s, t)$, 可通过在固定的 st 坐标(s^*, t^*)处的采样来获取二维切片 $I_{s^*, t^*}(u, v)$ 。切片 $I_{s^*, t^*}(u, v)$ (子孔径图像)可认为是由位于(s^*, t^*)处的相机捕获的照片。图 2 中的右侧部分为可视化样本子孔径图像。同样地,通过在固定的 uv 坐标 u^* 和 v^* 处进行采样来获取切片 $I_{u^*, v^*}(s, t)$ 。切片 $I_{u^*, v^*}(s, t)$ (通常被称为光场子视图)由每个点收集的来自不同视点的光线组成。通过收集两个空间维度或两个角度维度信息来获取上述二维切片。通过采集具有固定空间坐标 v 和角坐标 t (或 u 和 s)的光场样本,可以产生切片 $E_{v^*, t^*}(u, s)$ (或 $E_{u^*, s^*}(v, t)$), 得到的图称为极线平面图(EPI)。与针孔图像或光场子视图不同, EPI 包含空间和角度维度信息。考虑深度为 z 的具有朗伯表面的被观察物体上的某个点,当改变坐标 s 时,坐标 u 也将根据 $\Delta u = \frac{f}{z} \Delta s$ (f 是平面间的距离, Δs 为视角改变量)而改变,从而在 EPI 上形成一条线。具有不同深度的点将被可视化为 EPI 中具有不同斜率的线,参见图 2。相反, EPI 中线的斜率反映了由光场捕获的场景点的深度,这种特殊结构被广泛用于推断场

景几何。

2.2 光场渲染

四维光场 $L(u, v, s, t)$ 可通过由两个平行平面 st 和 uv 定义的相机捕获的一组视图定义。实际上, st 和 uv 空间也可以是球形的。与平行平面相比,球面在方向空间中可以更均匀地采样光线,并且到物体的距离固定。图 3 描绘了具有球面坐标的光场采样系统^[11], 物体位于球体的正中心, 球体表面分布有相机。只要球体完全包裹物体的凸包, 相机就可捕获物体的四维光场。

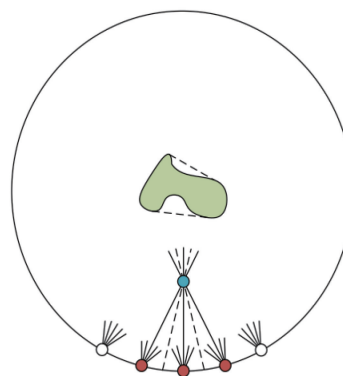


图 3 使用采集的球形空间得到的光场渲染虚拟视图

Fig. 3 Virtual light field rendering with the collected spherical space

使用足够密集的摄像机组,通过重采样和插值光线^[3]而不通过几何信息合成视图^[12],便可在球体表面的任何位置进行光场的虚拟渲染,甚至可在更接近物体的位置渲染。图 3 是在比采集视

点更靠近对象(绿色形状)的虚拟视点(位于蓝点处)处进行虚拟渲染的图示。对位于蓝色点附近但在其凸包外(中心对象周围的虚线)的蓝点的虚拟视图,因为光线辐射在自由空间中保持恒定,所以可在现有视点上重采样一些光线,如实线间的光线可通过重新采样位于红点处的摄像机捕获的光线获取。若光线不在被摄像机确切捕获的位置上,则需通过插值来推断(如图3中虚线所示的光线)。为渲染这样的光线,必须首先计算光线与两个坐标系的交点坐标,然后将最近的16个采样光线内插入虚拟光线。图4描绘了在四维光场的双平面表示下插值的情形,画出了四条用于插值的采样光线。上述方法称为光场渲染^[3]。

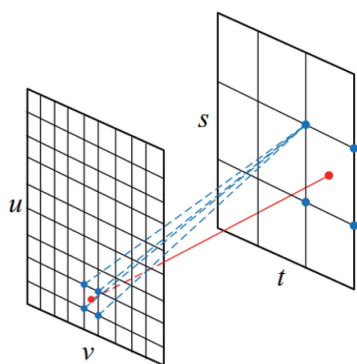


图4 四维光场中使用双平面表示插值的渲染情形

Fig. 4 Light field rendering with dual-plane interpolation in the four-dimensional light field

在光场渲染中,不足的光场采样将在虚拟视图中产生重影效应。然而,获得太多的光场采样是不切实际的。例如,Levoy等^[4]合成hallway光场时使用了四个光板,每个光板都涉及 64×32 张原始大小为1608 MB的图像。Chai等^[13]和Lin等^[14]研究了光场渲染所需的最小采样数,发现相邻视图之间的最大视差必须小于1 pixel,且视差越小,插值图像越清晰,于是得出结论:最小采样数是与相机分辨率、场景深度密切相关的值。然而,当场景的几何信息已知时,所需的采样数可大大减少,甚至不规则采样的光场也适合于新视图渲染。Buehler等^[15]提出了一个广义模型,它结合了使用不规则采样图像(即非结构化输入)的光场渲染和基于深度图像的渲染。

3 光场采集

本节重点介绍用于光场采集的现有设备和方法。传统相机通过对从各个方向到达每个像素的

光线进行积分来捕获光场在其传感器平面上的2D投影。相反,用于光场采集的装置或方法则通过区分光线的方向来测量光线的分布,以避免在同一像素位置对光强值的角度积分。但是,传感器只能测量空间两维度信息,为获得四维光场,需沿角度维度进行多次采样。现有的光场采集方法可分为三个类型^[8-10,16]:多传感器采集、时序采集和多路复用成像。

3.1 多传感器采集

多传感器采集方法需要用到分布在平面或球面上的图像传感器阵列,以同时捕获来自不同视点的光场样本^[3]。光场的空间尺寸(u 和 v)由传感器来确定,而角度尺寸(s 和 t)则由摄像机的数量及分布来确定。因此,多传感器采集方法通过采集图像的组合来记录4D光场。用于多传感器采集的设备如图5所示。

2002年,Yang等^[18]描述了一种使用 8×8 摄像机阵列进行动态光场捕获的设计,如图5(b)所示。为解决数据带宽的问题,他们采用了分布式渲染算法。2001年,Wilburn等^[20]使用6个CMOS图像传感器来记录同步的视频数据集,每个摄像头都有一个处理板,可实现MPEG2压缩,进行可扩展的采集。之后,他们进一步将系统扩展到125个摄像机,并用它采集每秒数万帧的视频^[17,21],该系统的一种配置如图5(a)所示。Zhang等^[22]提出了一个由 6×8 摄像头阵列组成的系统,该系统能够渲染新视图并重新配置摄像头的位置,以获得更高的渲染质量。Chan等^[23]使用 1×8 的摄像机阵列来采集和渲染“全光视频”的动态图像,该系统能够通过网络以15 frame/s的速率传输全光视频($256 \text{ pixel} \times 256 \text{ pixel}$ 空间分辨率)。Liu等^[24]还开发了一种基于宽带网络的 8×8 动态光场流系统。

典型的相机阵列系统体积大且昂贵,因此不适合大多数商业用途。由ViewPLUS公司制造的ProFusion 25^[25]包括装在一个小盒子里的 5×5 VGA摄像头阵列,能够以25 frame/s的速度采集光场视频,如图5(c)所示。Venkataraman等^[19]提出了一种超薄的单片相机阵列,他们将其称为PiCam(Pelican Imaging Camera-Array,如图5(d)所示)。PiCam是一个 4×4 相机阵列,每个相机均能够捕获 $1000 \text{ pixel} \times 750 \text{ pixel}$ 的图片。整个设备比硬币还小,可集成到手机中^[6]。对于微观场景,Lin等^[14]提出了一种基于相机阵列的光场显微镜设

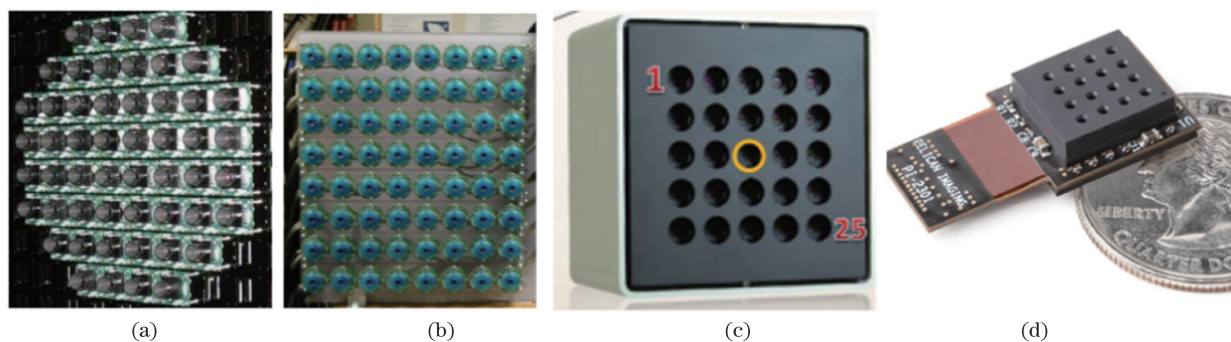


图 5 用于多传感器采集的设备。(a) Wilburn 等^[17]开发的 8×12 个摄像机组成的摄像机阵列系统；(b) Yang 等^[18]搭建的 8×8 摄像机阵列实现了实时光场渲染；(c) ViewPLUS 公司制造的大小仅为 90 mm×90 mm×60 mm 的光场相机阵列；(d) Venkataraman 等^[19]开发的 PiCam 是一个超薄单片 4×4 相机阵列，可集成到手机中

Fig. 5 Devices for multiple-sensor acquisition. (a) Camera array system composed of 8×12 cameras, developed by Wilburn et al.^[17]; (b) real-time light field rendering realized by the 8×8 camera array, constructed by Yang et al.^[18]; (c) light field camera array with size of 90 mm×90 mm×60 mm, manufactured by ViewPLUS; (d) PiCam, developed by Venkataraman et al.^[19], is a ultrathin single-chip 4×4 camera array and can be integrated into cell phone

置,用以测量通过样本体积的光分布。该系统如图 6 所示,包括一个传感器阵列,其分辨率为 1024 pixel×768 pixel,频率为 30 frame/s。

多传感器采集方法能够瞬时捕获光场,因此能

够记录光场时间序列。由于采用摄像机阵列,因此早期的多传感器系统的体积庞大且昂贵。然而,我们可以通过近期出现的更便宜和更便携的设计发现这种多传感器采集方法的潜力。

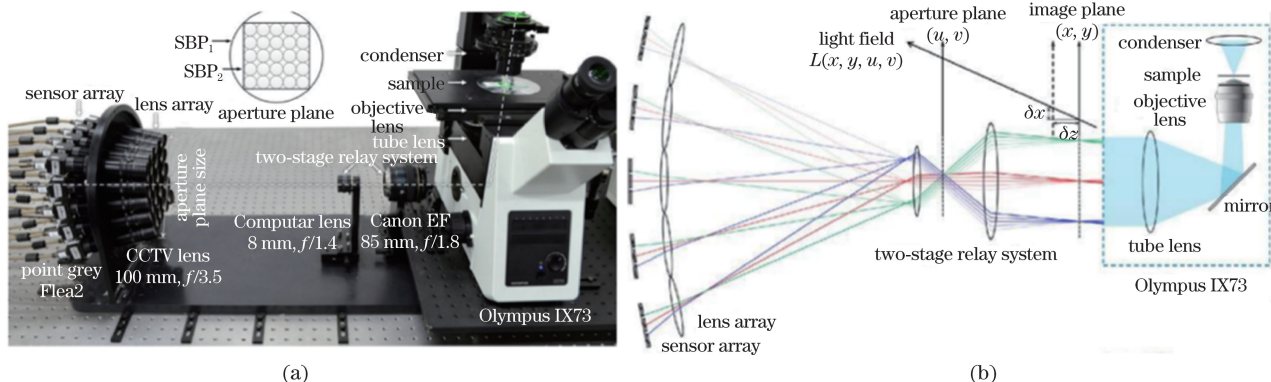


图 6 使用 5×5 相机阵列系统采集微观物体的(a)光场设置及(b)光路原理图^[22]

Fig. 6 (a) Light field setup and (b) optical path schematic for acquiring micro-object images with the 5×5 camera array system^[22]

3.2 时序采集

与多传感器方法相比,时序采集方法使用单图像传感器并通过多次曝光来捕获光场的多个样本。典型的方法是使用安装在机械台架上的传感器来测量不同位置的光场^[3,11]。例如,斯坦福大学计算机图形学实验室提出了两个龙门架系统^[26]:一个是计算机控制的龙门架,它有四个自由度——X 平移、Y 平移、俯仰和摇动;另一个是 Lego Mindstorms 龙门架,其电机具有旋转控制器,使摄像机能够沿着精确的路径移动,如图 7(a)

所示。Unger 等^[27]通过安装在电动线性平台上的单个摄像机(具有两个自由度)在 X 和 Y 方向的平移来捕获光场。图 7(b)显示了 Unger 等^[27]的系统。Kim 等^[28]提出了一种只有一个自由度的类似设计。该系统可用于捕捉室外具有高角分辨率的光场。为获得具有大视场的光场,Dansereau 等^[29]提出了一种紧凑的光学设计——将单中心透镜与微透镜阵列相结合(将在 3.3 节介绍),该系统安装在机械臂上,机械臂可围绕固定轴旋转,可以实现大视场角的捕获。

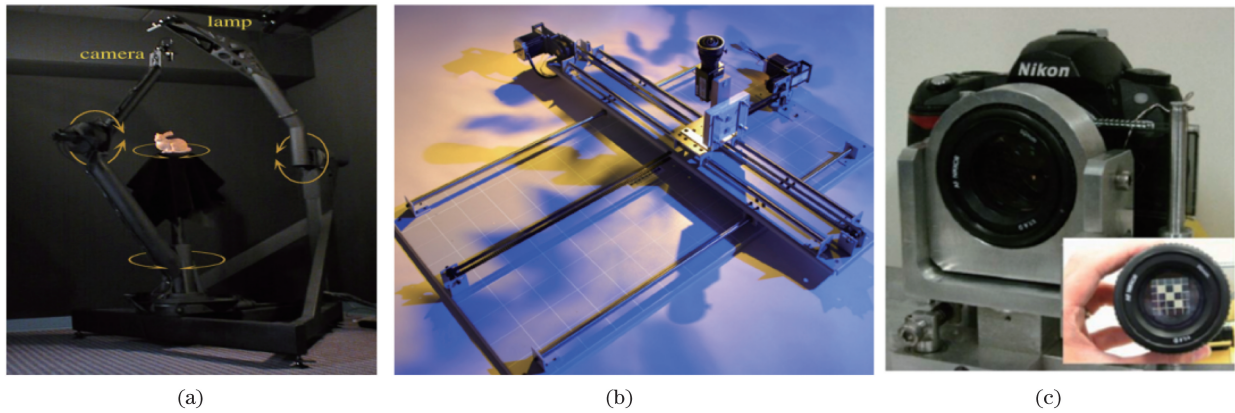


图 7 用于时序采集的设备。(a)斯坦福大学于 1996 年搭建的光场龙门架,可以实现静态光场的采集,通过物体、相机和光照的全自由度控制能实现高质量静态光场数据的采集^[3]; (b)Unger 等^[27]搭建的装有捕获四维光场的单个摄像机的电动线性平台; (c)Liang 等^[30]提出的可编程光圈,可通过多次曝光捕获光场

Fig. 7 Devices for timing sequence collection. (a) Light field framework, constructed by the Stanford University in 1996, to realize collection of the static light field by controlling all degrees of freedom including object, camera, and illumination^[3]; (b) electric linear stage, constructed by Unger et al.^[27], assembling a single camera for capturing the four-dimensional light field; (c) programmable aperture, proposed by Liang et al.^[30], for capturing the light field by multiple exposure

上述方法通过移动图像传感器在不同的视点采集图像,这不仅需要高精度的控制,而且耗时。幸运的是,人们提出了一些快速时序采集方法。Ihrke 等^[31]使用平面镜和高动态范围摄像机来记录光场,通过移动平面镜产生不同的视点供相机采集。Taguchi 等^[32]提出了一个由球镜和单摄像机组成的系统,该系统被称为“轴向光场”,该光场是通过沿着镜子的旋转轴移动相机来采集的。作为反射表面的替代,一个用于光场采集的可编程光圈被 Liang 等^[30]提出,如图 7(c)所示。孔径图案被不透明的纸张或液晶阵列编码,允许相机从特定角度捕获光线。通过使孔径图案与相机曝光同步,就可以使用全传感器分辨率采样光场。采集具有 5×5 角分辨率的光场共需要 25 个编码模板,每次曝光需要 10~20 ms。每个图像的空间分辨率均为 $3039 \text{ pixel} \times 2014 \text{ pixel}$,但都被下采样到 $640 \text{ pixel} \times 426 \text{ pixel}$,目的是实现高计算效率。

与多传感器系统相比,时序采集系统仅需要一个传感器,从而降低了整个系统的成本。此外,时序采集系统能够采集密集角度分辨率的光场,多传感器系统由于成本高而无法做到这一点。但是时序系统中的采集过程耗费时间,因此,它们仅适用于静态场景。

3.3 多路复用成像

这种方法旨在通过将角域复用到空间(或频率)域中来将 4D 光场编码到 2D 传感器平面。它允许使用单个图像传感器进行动态光场的捕获,但需在

空间和角度分辨率之间进行权衡(即若想在空间域中密集采样,需在角域中稀疏采样,反之则反)。多路复用成像可进一步分为空间复用和频率复用。

3.3.1 空间复用

空间复用中,采用传感器采集来自光场的不同 2D 采样切片交错排列的元素图。大多数空间复用方法使用安装在图像传感器上的微透镜阵列或小透镜阵列来实现。有趣的是,这是光场成像的最早方法之一:1908 年,Lippmann^[33]在他的“整体摄影”的开发中使用了这种方法。

1992 年,Adelson 等^[34]描述了一种用于“单镜头立体”的新型相机,它包含一个主镜头和位于传感器之前的透镜阵列。该设备可以捕获 5×5 角分辨率的光场,每张图片大小为 $100 \text{ pixel} \times 100 \text{ pixel}$ 。Ng 等^[35]展示了在传感器和主镜头之间插入 296×296 小透镜阵列的手持式光场相机,该设备的操作类似于传统相机,可以捕获 14×14 角分辨率光场,如图 8(a)所示。

上述空间复用装置属于“全光相机 1.0”类型;其中,每个微透镜采集场景中的一个位置。属于“全光相机 1.0”类型的两款商用相机是 Raytrix^[5]和 Lytro^[36]相机,它们分别适用于工业消费者和个人消费者。Wei 等^[37]提出了一种新颖的设计,将采样不规则性和镜头像差引入光场相机,用以提高其质量和可用性。“全光相机 1.0”中的每个微透镜捕获光线的角度分布,通过汇集每个子视图在

相同坐标下的像素,可以获得位于特定视点的图像。另一种版本使用多个微透镜来采集场景中的相同位置,称为“全光相机 2.0”,其角度信息分布在每个微透镜上^[38]。

一些光场采集方法也使用滤波器和镜面阵列而不使用微透镜阵列。Horstmeyer 等^[39]使用置于主透镜光瞳平面中的多个滤波器和传感器平面上的针孔阵列来采集光场,交换和重新配置滤波器,该方法具有高度的灵活性。Levin 等^[40]在相机镜头的光圈内插入一个有图案的掩模,形成了一个编码光圈。

利用图像的统计模型,从改进的相机拍摄的单个照片中恢复深度信息和全焦点图像,即光场信息,如图 8(d)所示。同样,Manakov 等^[41]开发了一个可重新配置的附挂,标准相机安装附挂后可进行全光成像,即进行多光谱、高动态范围光场成像。以类似于 3.2 节中描述的“轴向光场”采集方式来采集由球面镜反射的光场,Unger 等^[27]和 Lanman 等^[42]提出了用一系列球面镜来捕获光场的方法。略有不同的是,他们的设计使用了多个镜子,因此将时序采集转换为空间多路复用,如图 8(c)所示。

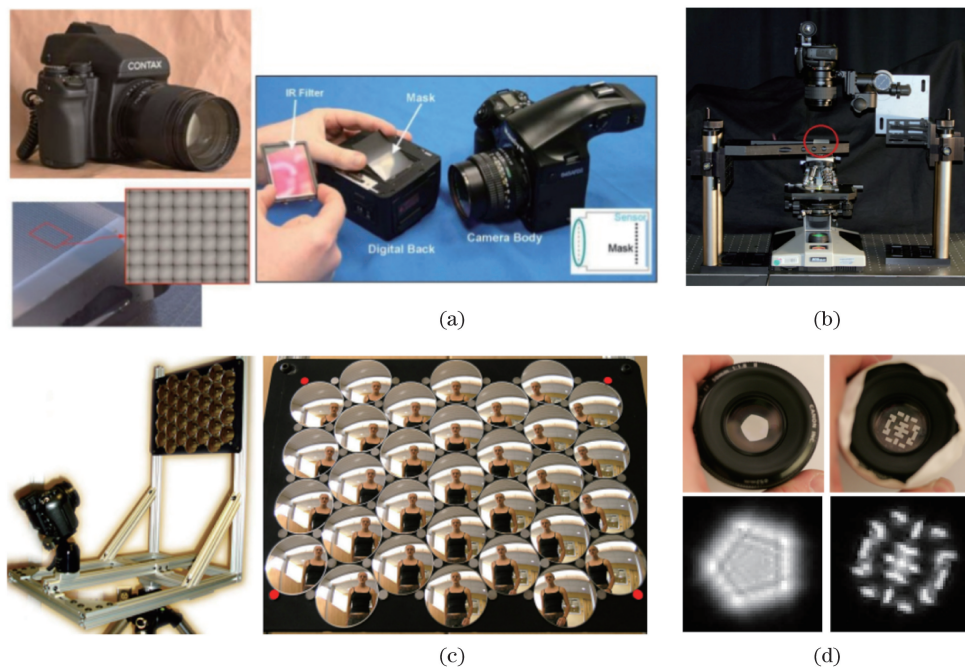


图 8 用于多路成像的装置。(a)微透镜光场相机及其原理^[35];(b)斯坦福大学开发的基于微透镜阵列原理的光场显微镜^[7];
(c) Lanman 等^[42]开发的多反射球光场采集装置;(d)Levin 等^[40]开发的基于掩模的光场复用采集方法
Fig. 8 Devices for multi-channel imaging. (a) Microlens light field camera and its principle^[35]; (b) light field microscope based on the microlens array, developed by the Stanford University^[7]; (c) multiple reflective sphere light field acquisition device, developed by Lanman et al.^[42]; (d) light field multiplexing acquisition method based on masks, developed by Levin et al.^[40]

除测量宏观物体或场景之外,还可使用空间多路复用来捕捉微尺度的样本^[7,43-45]。Levoy 等^[7,43]将微透镜阵列插入传统显微镜的光路中,以在单个传感器平面上记录生物标本光场,如图 8(b)所示。与基于相机阵列的光场显微镜系统相比^[46],Levoy 的系统采集到了具有高角分辨率(17×17)的光场,但空间分辨率较低。Mignard-Debise 等^[47]使用消费者光场相机 Lytro 和光学匹配系统开发了一种光场显微镜,获得了大约 $6.25 \mu\text{m}$ 的最大空间分辨率。

空间复用是采集光场最广泛使用的方法,可基于单次曝光采集的光场形成单图像。然而,空间多

路复用方法存在其固有的问题,即在图像传感器处的角度和空间分辨率之间的折中。为解决这个问题,Georgiev 等^[48]采用牺牲角分辨率的方式获得了更高的空间分辨率。他们展示了一个 4×5 阵列的镜头和切成正方形的棱镜,并将它们放置传统相机前面,以采集低角度分辨率的光场。然后,应用插值技术来合成密集角分辨率光场。许多研究人员专注于空间/角度范围内的超分辨率光场,请参阅第 4 节有关超分辨率的更多详细信息。

3.3.2 频率复用

与在传感器平面上交错排列的 2D 光场切片的空间复用方法不同,频率复用方法将光场的不同 2D

切片编码到不同的频带中。通常,频率复用方法使用调制掩码来实现傅里叶域中的某种属性^[16]。

Veeraraghavan 等^[49]描述了在传统图像传感器光路中使用衰减掩模的 4D 光场采集的理论框架,他们称之为“斑驳摄影”。图案化的掩模不是混合光线,而是在图像传感器上对光线进行衰减和编码处理。将傅里叶变换图像重新排列为 4D 数据,然后应用傅里叶逆变换来恢复光场。如 Lanman 等^[50]所述,图案化掩模相当于对高角度采样率的针孔阵列的截断傅里叶级数近似。Amit 等^[51]进一步描述了两种用于光场成像的独立结构,其允许在角度或空间维度上进行压缩光场成像。他们的设计是在主透镜后面采用振幅掩模进行角度压缩光场成像,并在每个微透镜上放置振幅掩模,用于空间压缩光场成像。Antipa 等^[52]使用漫射板编码多路复用的空间和角度信息,与衰减掩模相比,漫射板允许更大的光通量通过,且提供了用于频率复用的廉价的、灵活的装置。Pégard 等^[53]采用压缩光场成像技术在微观尺度上对标本进行采样,而 Cohen 等^[54]则采用波前编码技术在深度上产生更均匀的光场分布,以提高光场显微镜的性能。利用学习的技术,Marwah 等^[55]描述了一种基于衰减掩模的方案,并使用过完备字典来从编码的 2D 投影中重建光场,如图 8(c)所示。Wang 等^[56]使用随机卷积 CMOS 传感器进一步改善了重建光场的质量,该传感器能够通过相关性保留更多信息。

尽管频率和空间复用是两种不同的复用成像方法,但它们在复用模式方面有密切关系。Ihrke 等^[16]的研究表明,可通过空间插值来重建在傅里叶空间使用的模板,且还可以使用傅里叶空间算法重建空间模板。此外,频率复用的傅里叶重建相当于使用 Sinc 滤波器核进行空间重构,且可通过空间滤波器(例如 cubic 插值)或边缘保留滤波器显著消除产生的环状伪像。

4 面向空间高分辨的计算光场成像

如前所述,光场成像旨在对真实物理世界的内容进行感知采样,其感知能力与采样效率可用空间带宽积(SBP)进行诠释。空间带宽积从信息学的角度描述了一个成像系统的信息容量,是对其视场大小与空间分辨率的联合统计,从本质上决定了成像系统的性能。根据香农采样定理,对于成像系统这样一个空间、频域同时受限的信号系统,空间带宽积正比于其视场面积与空间频域面积的乘积。大空间带宽积的光学系统具备更高的信息自由度和更大的数据通量,但光学像差无疑严重制约了这一系统要素,导致对真实世界的感知限定在有限的空间带宽积范围内。无论是现有的各类宏观成像的单反镜头还是显微成像领域的物镜,它们的有效空间带宽积都十分有限,往往仅在千万像素级别,远远无法达到大场景 10 亿像素的需求。本章节重点关注面向空间高分辨的计算光场成像。

单像传感器成像:在实时帧率采集下,现有成像系统像传感器的空间分辨率难以突破亿像素级,宽视场范围和高细节分辨率之间存在难以调和的矛盾,视觉成像难以兼顾成像范围与成像分辨率。为实现超高分辨率成像,微软研究院 Kopf 等^[57]于 2007 年首先提出用一个电机控制的云台和配备长焦镜头的单反相机来采集静态 10 亿像素图像的方案。此方案成本低,算法成熟,其相应的商业产品如 Gigapan^[58]已被普遍应用(图 9)。受云台扫描效率的限制,Gigapan 只能采集静态图像,其本质上是通过牺牲时间换取了空间分辨率。随后,Pirk 等^[59]在此基础上增加了一个相机来采集视频片段,并将这些视频片段嵌入到十亿像素图像中。但该工作仅仅为 10 亿像素图像背景添加了部分动态内容,并不能为视觉计算提供源自采集端的动态数据支撑。除此之外,必须用相同的设备同时采集背景静态图像和嵌



图 9 10 亿像素图像采集设备 Gigapan 及拼接所得的 1.5 亿像素全景图^[58]

Fig. 9 Giga-pixel image acquisition device Gigapan and stitched panoramas with 1.5×10^8 pixels^[58]

人的动态视频,拼接结果依靠离线算法生成。

结构化像传感器阵列:哥伦比亚大学的 Nayar 教授课题组^[60]在 2011 年提出了基于球透镜分级成像的 10 亿像素相机的原理与系统,并从理论上对这种设计进行了详细分析^[61]。如图 10(a)所示,物体通过前端球透镜进行一次成像后,在后端用

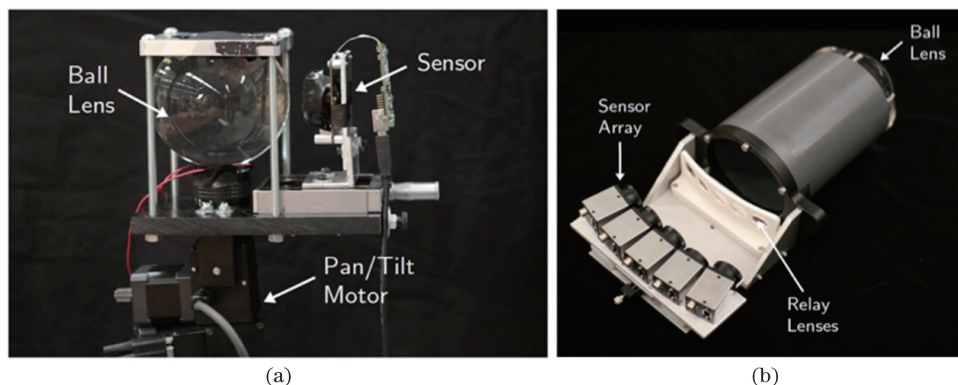


图 10 基于分级成像的 10 亿像素相机原型机^[60]

Fig. 10 Prototype of gigapixel multiscale camera^[60]

杜克大学的 Brady 教授课题组^[62]于 2012 年在《Nature》上发表了关于通过前端球透镜进行一次成像后,后端通过分视场相机阵列(约 98 个)进行采集的 AWARE-2 相机阵列的研究论文。如图 11 所示,AWARE-2 相机阵列采用的是多重光学镜头的设计,受益于球透镜的成像单一光心特性,该设计避免了分视场相机之间的视差,大大降低了图像拼接

分视场相机进行采集。受益于球透镜的单光心特点,这种设计避免了不同相机之间的视差。尽管通过像传感器阵列的拼接可初步验证动态 10 亿像素成像的可行性,但其原型系统仅包含了 5 个像传感器,如图 10(b)所示,并没有实现 10 亿像素级成像。

的难度,可达到每秒几帧的 10 亿像素采集速度。尽管通过像传感器阵列的拼接可初步验证动态 10 亿像素成像的可行性,但 AWARE-2 相机阵列需要特制的镜头,而且相机阵列的装调和标定极其严格、复杂,其中大规模分视场像传感器的引入对实际安装、标定和拼接均带来了极大的技术挑战,实现成本高、难度大且灵活性低^[63-64]。



图 11 AWARE-2 相机阵列及成像结果^[62]

Fig. 11 AWARE-2 camera array and imaging results^[62]

虽然 AWARE-2 相机阵列实现了 10 亿像素成像,但是受限于目前的传感器技术,球形镜头和后面的传感器阵列之间需要有足够大的距离才能保证传感器相互之间有重叠区域,导致采用该方案设计的相机阵列存在体积偏大的问题。因此,Pang 等^[65]提出了并联单光心相机阵列的方案,即通过组合多个单光心的球形分级相机阵列来得到一个像素数量基本相同,但体积更小的相机阵列,如图 12 所示。这种新的

设计能够将相机阵列的体积缩小到原来的 1/10,但是由于不同组的相机(图 12 中不同颜色的圆点)之间存在视差,图像拼接的效果和难度会大大增加。

斯坦福大学的 Dansereau 等^[29,66]将上述单光心球形阵列相机与光场相机结合,搭建了宽视角光场相机原型。如图 13 所示,该设计有效地提升了现有光场相机的视场角,但受限于微透镜阵列基线的大小,该原型相机只能对 5 m 内的场景进行深度估

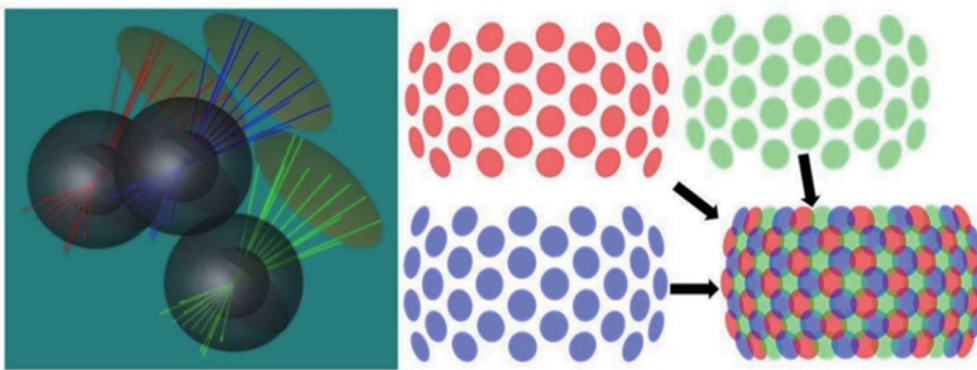


图 12 由三个单光心球形分级相机阵列组成的并联单光心相机阵列^[65]

Fig. 12 In-parallel single-photocenter camera array composed of three single-photocenter spherical multiscale camera arrays^[65]



图 13 基于单光心球形阵列相机与光场相机的宽视角光场成像

Fig. 13 Wide-field-of-view light field imaging based on single-photocenter spherical arrayed camera and light field camera design, application is very limited.

异构像传感器阵列: Wang 等^[67]提出了高质量混合光场合成系统,设计出了可改造单反相机为光场采集相机的外置组件,该组件可通过设定进行同步采集。该装置充分利用了单反相机成像质量高和普通相机成本低的特点,通过大范围环状采集,既克服了相机阵列不便携的劣势,也克服了微镜头阵列空间分辨率普遍下降的问题,实现了高空间、高角度分

辨率及低成本的光场采集功能。

基于单反相机的异构光场相机阵列系统如图 14 所示。左图为单反相机附带光场改造组件,异构相机平台由中间水平放置的高分辨率单反相机和环绕四周的 8 个水平放置的商用低分辨率 USB 相机构成。其基本设计思路分为两步:首先由 9 个相机采集一个场景图像,包括中间的单反相机高分辨率图像和周围的 8 个低分辨率图像;之后以高分辨

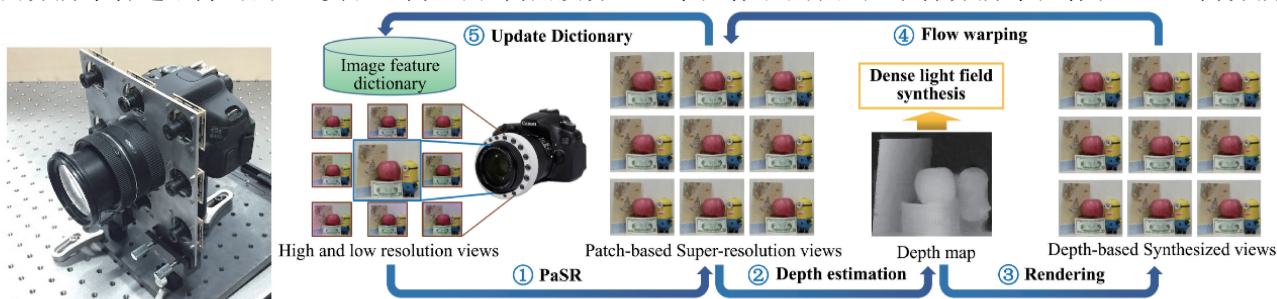


图 14 混合光场高分辨率计算成像系统

Fig. 14 Hybrid light field high-resolution computational imaging system

率图像为参考,提升周围 8 幅图像的质量,获得超分辨率图像,用于密集光场的合成。基于硬件采集平台,Chen 等^[11]进一步提出了基于图像块和深度的迭代式合成算法(iPADS),计算框架如右图所示。通过反复迭代可以获得深度信息,将深度信息用于光场合成,可以提高合成光场的质量,算法稳定性和运行时间均可满足实际应用。

上述结构化像传感器阵列架构本质上与单像传感器感知遵从同样的假设,即均匀化采样、信号密度同分布,相邻像传感器间遵从视场相互重叠的严格限制。均匀化感知势必面临时空带宽积的瓶颈,且光学成像装置面临物理体积随数据通量线性增长的制约,维度扩展性严重受限。

非结构化异构像传感器阵列:考虑到现实场景中

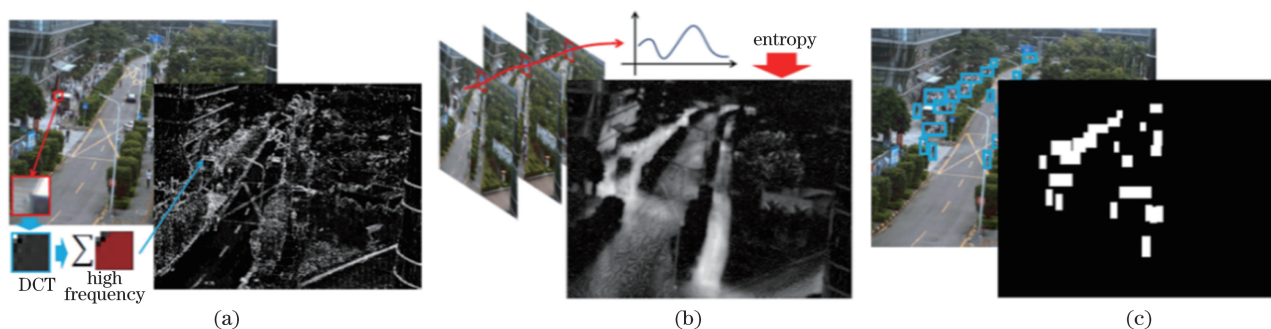


图 15 一个自然场景的数据稠密度示例。(a)空间上,整个图像分为 8×8 块,使用高频系数之和表示空间信息的密度;(b)时间上,数据密度由沿时间轴的每个像素信号值的熵表示;(c)语义数据密度由检测到的人、车辆边界框代表的高级信息分布表示

Fig. 15 Data density example of a natural scenario. (a) The image is divided into 8×8 pieces spatially and the sum of high-frequency coefficients is used to represent the spatial information density; (b) data density is represented temporally by the entropy of each pixel signal along the time axis; (c) semantic data density is represented by the high-level information distribution labeled with the tested people boundaries and vehicle boundaries

Yuan 等^[68]另辟蹊径,提出了非结构化光场感知算法,构建了“多层异构-同层自适应”的树形感知模型。其中,“多层异构”一改传统结构化阵列像传感器固有的图连接感知模型,解除了光学成像装置物理体积随数据通量线性增长(Scaling Law)的制约,突破了现有多维多尺度感知对庞大复杂硬件系统的依赖。与此同时,“同层非结构”突破了现有像传感器均匀化采样、信号同分布的假设,相邻像传感器间无需受视场相互重叠的严格限制,各像传感器面向场景内容独立感知、并行计算,突破了传统均匀化感知所面临的时空带宽积的瓶颈,实现了像传感器阵列硬件复杂度、数据感知量(信息容量)、维度扩展性的联合优化。

图 16(a)所示为传统结构化像传感器阵列,其成像感知遵从均匀化采样机制,每个像传感器节点通过

包含的信息大多具备不规则和非凸特性,传统成像感知所基于的均匀采样假设显然并不完全成立。图 15 为典型自然场景下的数据密度分布示例。在空间维度上,真实场景采样所得的数据密度与图像视频的高频分量密切相关;在时间维度上,具有移动物体如“人”“车辆”的区域相比于静态区域具有更大的数据密度;在语义维度上,高层信号呈现更为明显的稀疏、非均匀特性。结合空间、时间、语义等维度的数据特性,自然场景中重要内容的占比为 $10\% \sim 30\%$,且呈现非均匀分布的特性。尽管常用的先成像、后压缩的感知模式可在一定程度上降低数据的冗余度,但究其根本,依然受制于均匀采样假设,因此难以实现场景内容的自适应性和可扩展性。

部分重叠区域连接到其相邻的像传感器节点,所感知的数据遵循“图”连接结构。在均匀化采样的假设下,将拼接得到的全景图像/视频投影为矩阵形式并进行数据的表征与存储。此外,各个像传感器均遵从双向图连接,单像传感器的调整势必会影响其他所有像传感器。图形结构的另一个问题是,由于其位姿估计严重依赖重叠区域的全局优化,随着像传感器数量增加,复杂度急剧增加,鲁棒性显著降低,可扩展性严重受限。与此不同,文献[68]提出的非结构化成像感知遵从非均匀采样机制,各像传感器面向场景内容独立感知、并行计算,以可分离的树形结构方式连接。如图 16(b)所示,该树形结构涉及分层架构,赋予层间异构互联、层内非结构化的特性。其中,分层架构的父层像传感器与子层像传感器之间自然重叠,同层内的各像传感器之间则完全非结构化,无需遵守任

何重叠限制,从本质上赋予像传感器阵列极高的灵活性和可扩展性。在此基础上,各个局部像传感器的调整

(移除/添加)不会影响其他像传感器,亦不会对多像传感器位姿的优化带来复杂度剧增和鲁棒性降低的问题。

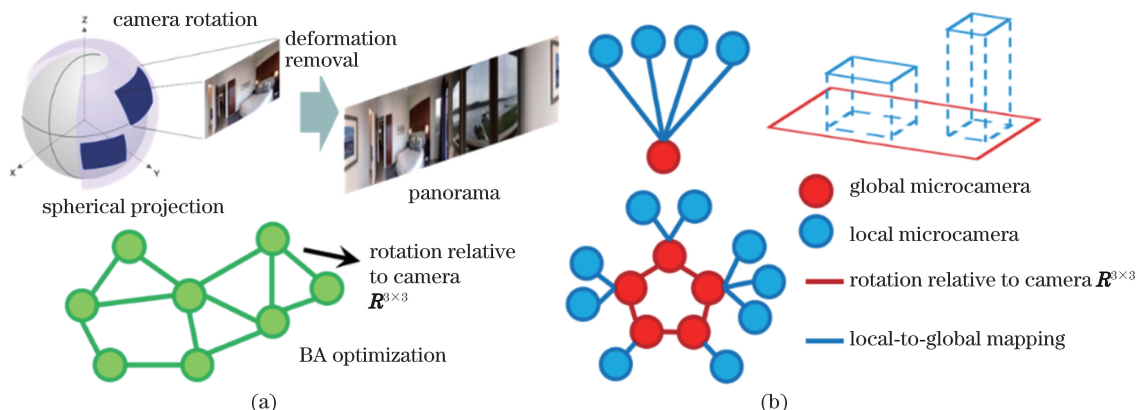


图 16 像传感器阵列。(a)传统结构化像传感器阵列;(b)非结构化像传感器阵列

Fig. 16 Image sensor array. (a) Conventional structured image sensor array; (b) unstructured image sensor array

图 17 展示的非结构化异构像传感器阵列由一个全局像传感器(红色)和 19 个局部像传感器(蓝色)组成。所有的像传感器均为 4000 pixel×3000 pixel 分辨率的 FLIR FL3-U3-120S3C-C 像传感器,其中全局像传感器作为辅助像传感器,配备了 16 mm 镜头(chiopt FA1615A,24°水平视场角,1/2.3" CMOS),为 10 亿像素计算成像提供低全局内容参考。局部像传感器配

备了变焦 25~135 mm(Computar E5Z2518C-MP)镜头,根据不同的物距来选择合适的焦距。图中两个局部像传感器的焦距分别为 72 mm 和 135 mm,分别覆盖近处和远处区域。传统的结构化像传感器阵列要求像传感器彼此重叠,而非结构化像传感器阵列受益于全局像传感器的指导/参考作用,局部像传感器之间无需视角重叠,实现了高灵活性和扩展性。

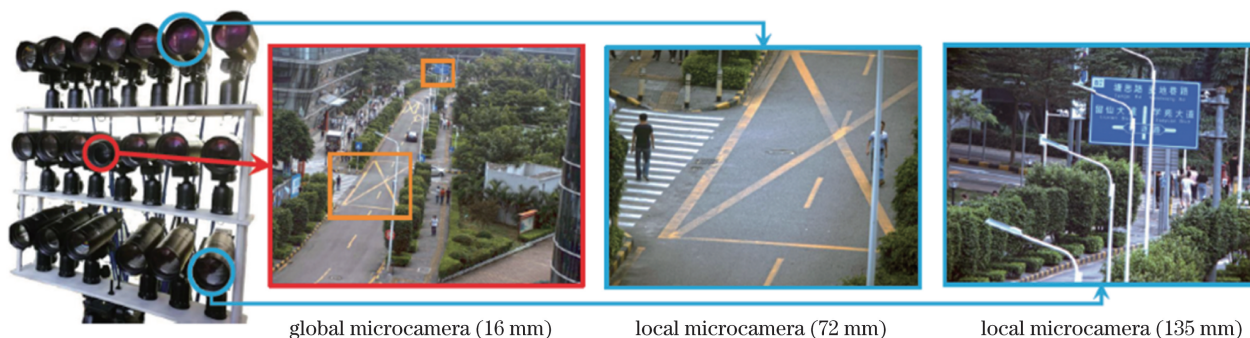


图 17 非结构化异构像传感器阵列^[68]

Fig. 17 Unstructured heterogeneous image sensor array^[68]

非结构化异构像传感器阵列本质上是多个极低成本相机,以多尺度方式灵活地组合,基于先进的计算成像思想,将硬件的成本以算法替代,实现实用且性价比高的 10 亿像素视频相机。基于该原理的相机系统通过分视场采集突破了空间带宽积的限制,同时通过引入全局尺度相机克服了分视场相机的像差与畸变,并通过跨尺度视频融合算法规避了复杂的相机系统的装配和标定,同时回避了紧密光学部件的设计和制造需求,极大地降低了系统的成本。但是,目前成像性能远未发挥多尺度成像带来的益处,多尺度成像原理和跨尺度映射与融合机制仍未被揭示,系统设计(包括全局低

分辨相机和局部高分辨相机各自的设计)的灵活性仍未被充分挖掘,全局低分辨相机与局部高分辨相机的互补性未被充分优化利用,新型、高效、高性能的非结构化像传感器阵列的研究具有广阔的空间。

5 面向视角高分辨的光场成像

现有光场采集机制面临空间分辨率与角度分辨率的固有矛盾。一方面,采用微透镜阵列原理的光场采集设备(如 Lytro Illume^[4]与 Raytrix^[5])通过将各个角度的光线直接分散到成像单元上,来获取高角度分辨率的光场图像,但受到成像单

元分辨率的限制,其空间分辨率往往远小于如今比较通用的成像设备(单反相机、手机等);另一方面,相机阵列系统虽能够通过改变每个相机的分辨率来控制光场空间分辨率的大小,但受到相机物理空间以及系统复杂度、整体成本等因素的限制,其角度分辨率远远无法满足实时光场渲染的要求^[3]。为了同时兼顾光场采集设备的轻量化与

高质量光场信息的获取,需要通过计算成像的方式对稀疏采样的光场进行重建。国内外目前主要采用稀疏化角度维度来采集具有高空间分辨率的光场,如图 18 所示。根据对场景几何结构的依赖性,光场超分辨率方法可大致分为以下两类:基于场景深度或几何的视点合成方法和隐式场景深度的光场重建方法。

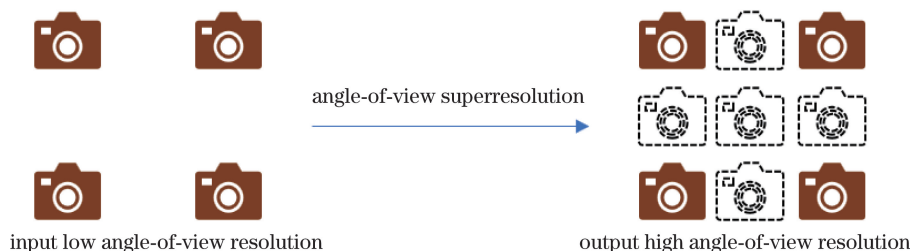


图 18 采用稀疏化角度维度采样方式进行高分辨率光场重建

Fig. 18 Reconstructing high-resolution light field by sparse angle dimension sampling

5.1 基于场景深度或几何的视点合成方法

典型的基于深度的视点合成方法流程如图 19 所示。该方法通常需要先对场景进行深度估计,例如:Wanner 等^[69]在 EPI 域下利用结构张量方向估计方法来获取每个像素的视差;Georgiev 等^[48]采用基于分割的光流方法来计算图像和两个相邻视点之间的光流;Zhang 等^[70]采用基于频谱相位的视差估计方法对两个光场子视点图像进行深度计算。在获取深度信息后,该类方法通常将已有的视点图像投影到新的视点下,最后采用某种特定方式将这些投

影后的视图进行合成。Pearson 等^[71]将子视点图像中的每个像素按深度信息分配给特定图层,然后使用概率插值方法合成了任意新视点的图像。在此基础上,Penner 等^[72]将深度信息在估计过程中的不确定性以及遮挡关系考虑在内,提出了一种柔性视点合成方法。Wanner 等^[69]则将视图合成问题转化为最小化能量方程的优化问题。Chaurasia 等^[73]提出了一种使用超像素技术的新变形方案,然后将变形视图用于由相机方向和深度可靠性确定的权重融合。

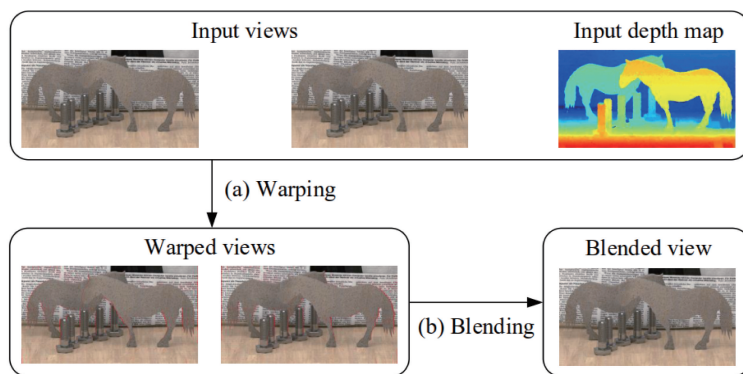


图 19 典型的基于深度图像视图合成的流程图

Fig. 19 Flow chart of typical view blending based on depth map

传统基于深度的视图合成方法严重依赖于深度估计的精度,对无纹理区域和遮挡区域的干扰较为敏感。此外,这些方法通常关注深度估计的精度而非合成视点的质量。近年来,随着卷积神经网络(CNN)在计算机视觉和图形学领域的发展,一些利用 CNN 优化合成视图质量的研究取得了良好的应用效果。Kalantari 等^[74]使用两个序列 CNN 来分

别进行深度估计以及视点合成,并采用端到端的优化方式最小化合成视图与真实训练标签图像之间的误差来训练两个网络(如图 20 所示)。但该方法需要对每个视点进行新的深度估计与视点合成,在所需光场视角分辨率较高时,计算速度会大幅下降。

Zhou 等^[75]提出了一种多平面图的光场表达方法,并将其用于快速合成新视点图像。与传统图像

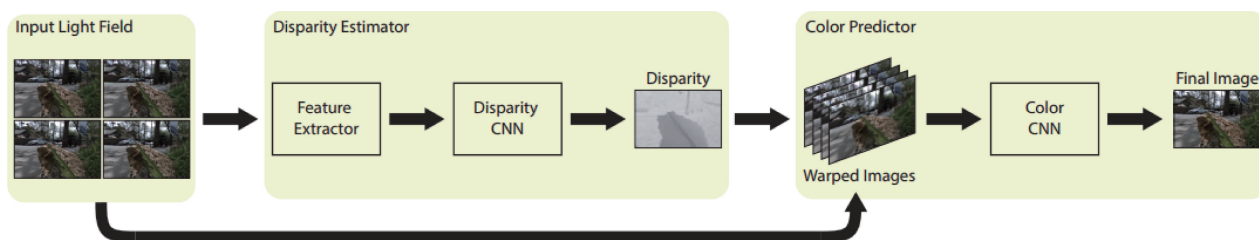


图 20 通过两个序列 CNN 分别进行深度估计与视点合成

Fig. 20 Disparity estimation and color prediction via two CNNs

分层^[71]不同的是,该方法利用 CNN 将图像分解为多个彩色图像层以及透明度层,其中透明度层用于对不同图层进行加权合成。在此基础上,Mildenhall

等^[76]将稀疏视角采样光场中的每个视点分解为多平面图,并利用其对局部光场进行合成,最终形成一个宽视角的光场,如图 21 所示。

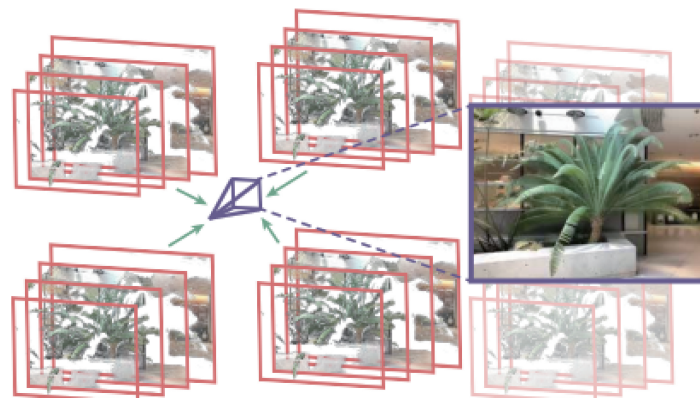


图 21 稀疏视角采样光场的多平面图分解方式与视点合成

Fig. 21 Multi-plane decomposition and color prediction of light field by sparse angle-of-view sampling

5.2 规避显式场景深度的光场重建方法

基于场景深度或几何的视点合成方法由于依赖深度估计的精度,往往在遮挡区域、透明/半透明等非朗伯区域产生很大的缺陷,如图 22 所示。而规避显式场景深度的光场重建方法通常基于全光函数的采样和连续重建,将光场视为一种高维信号,采用信号的采样与恢复策略对其进行重建,其难点在于对欠采样信号的抗混叠处理上。经典的光场重建理论^[4,77]通常从傅里叶域入手,根据原频谱及其复制频谱产生的混叠设计一定的抗混叠

重建滤波器以实现光场的重建,如图 23 所示。Shi 等^[78]将光场重建问题转化为连续傅里叶域下的稀疏优化问题,通过对矩形边框及其对角线上的视角进行采样,然后进行光场的重建,这种做法对采集方式有较为严格的要求。Vagharshakyan 等^[79-80]在傅里叶域下利用一系列剪切波构成一个复合滤波器,从而严格包裹原频谱信号,并过滤由欠采样导致的混叠频谱信号。然而这种重建方式依赖于初始的深度信息,用于引导复合滤波器的设计。

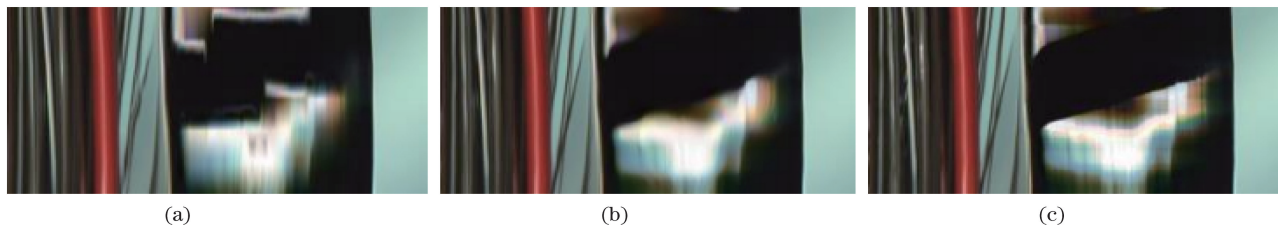


图 22 规避显式场景深度的光场重建方法在非朗伯区域具有优异的效果。(a) 基于场景深度估计的视点合成结果(EPI); (b) 规避显式场景深度的重建 EPI;(c) 目标 EPI

Fig. 22 Light field reconstruction, circumventing explicit disparity, has excellent effect on the non-Lambert area. (a) EPI based on disparity estimation; (b) reconstructed EPI circumventing explicit disparity; (c) target EPI

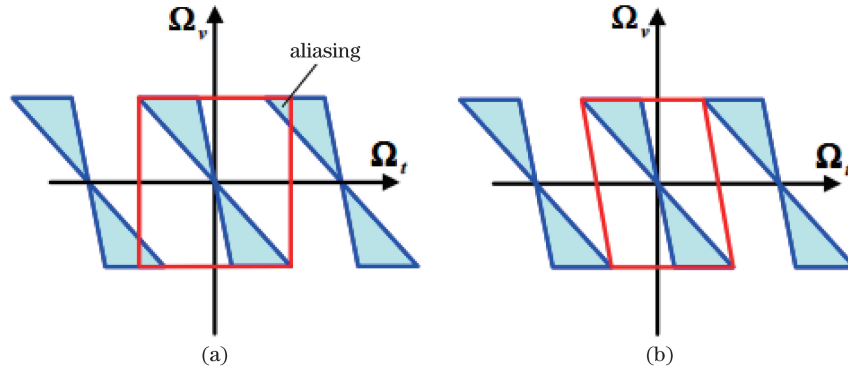


图 23 经典频域光场重建方法。(a)光场欠采样会产生典型的混叠现象；
(b)通过引入抗混叠重建滤波器(如红色四边形所示)重建光场

Fig. 23 Classical frequency-domain light field reconstruction. (a) Typical aliasing resulting from light field under-sampling; (b) light field reconstruction by using anti-aliasing filter (shown as the red quadrangle)

近年来,国内外研究人员也提出了一些基于深度神经网络的方法,并采用这些方法在不进行显式深度估计的前提下对光场进行重建。Yoon 等^[81]提出了一种基于 CNN 的方法,该方法在垂直(水平)角度维度中使用两个相邻视图来重建单个中间视图。然而,该方法未充分利用整个角度信息,且只能进行固定超分辨率的重建。而 Yeung 等^[82]则直接将四维光场数据送入高维卷积神经网络中,并采用角度与空间维度混合编码的方式实现各维度之间的信息融合。此外,为解决高维神经网络所带来的大

规模网络参数的问题,该方法采用伪四维卷积核,即将卷积核两个维度的大小置为 1,并将其作为主要的卷积层。Wu 等^[83]利用 EPI 清晰的纹理结构,将角度超分辨率建模为基于 CNN 的 EPI 角度信息恢复问题,如图 24 所示。在将 EPI 输入到网络之前,通过提取空间维度的低频分量来对抗 EPI 角度维度欠采样时的混叠效应,并采用非盲解卷积方法对空间维度进行恢复。该方法对复杂结构以及不同采集形式获得的光场图像的重建效果均很优异,如图 25 所示。

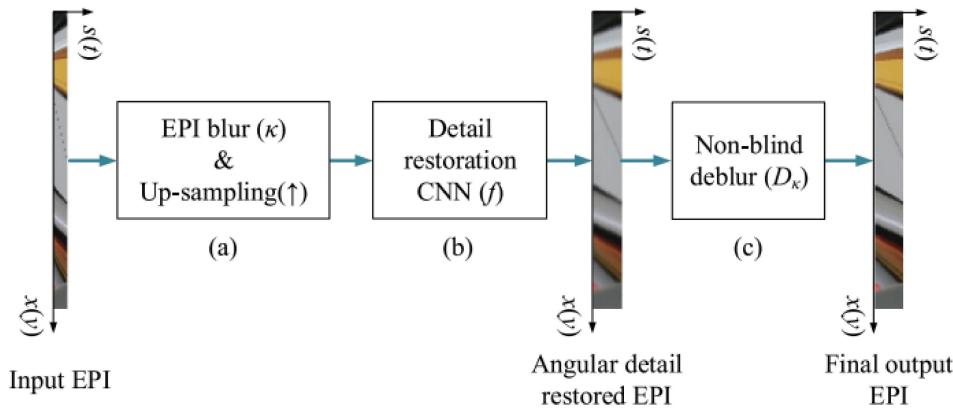


图 24 利用极线平面图(EPI)进行光场重建。(a)输入深度神经网络前,提取 EPI 空间维度的低频分量进行抗混叠处理;
(b)利用深度神经网络进行角度维度的信息恢复;(c)利用非盲解卷积恢复 EPI 空间维度信息

Fig. 24 Light field reconstruction via EPI. (a) Extract low-frequency components in the EPI spatial dimension for anti-aliasing before input into the deep neural network; (b) recover information in the angle dimension through the deep neural network; (c) recover information in the EPI spatial dimension through non-blind deconvolution

6 总结与展望

本文对计算光场成像装置与算法的研究进行概述和讨论,内容涵盖了光场表示和理论、光场信号采集、空间与视角维度重建等。同时,结合现有的研究

领域以及光场独特的属性,对全光函数新理解、动态光谱光场成像及光电计算光场视觉进行了展望。

6.1 全光函数新理解

七维全光函数用于记录某一时刻 t , 在三维空间任意位置 (x, y, z) 沿着方向 (θ, φ) 观察到频率为

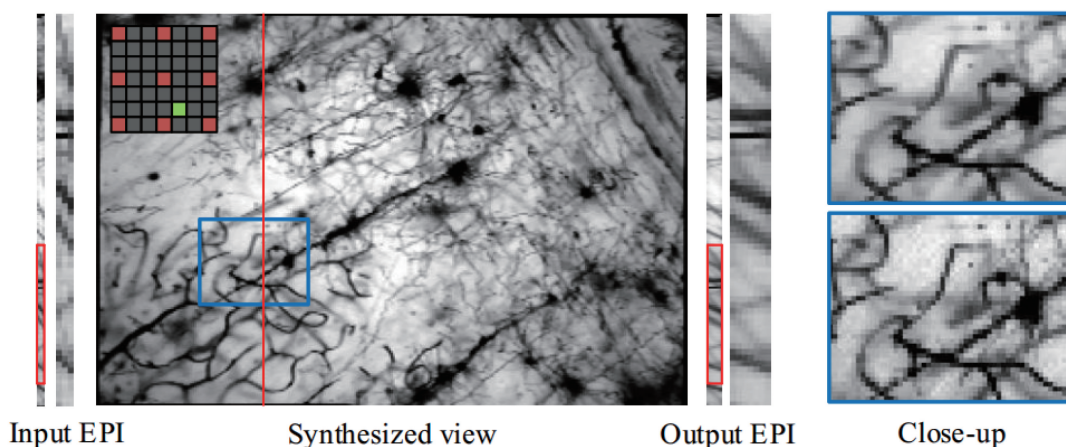


图 25 显微场景下的光场重建,输入为 3×3 视角分辨率的光场,输出为 7×7 视角分辨率的光场。

右上为 Wu 等^[83]所提方法的输出结果,右下为目标结果

Fig. 25 Light field reconstruction in the microscopic scenario with input and output light field angle-of-view resolutions of 3×3 and 7×7 , respectively. The top right shows the output reported by Wu et al.^[83] and the bottom right shows the target result

λ 、强度为 $|f(x, y, z, \theta, \varphi, \lambda, t)|$ 的光线。经典的成像模型是基于七维全光函数的二维投影子空间采样。通常,传统相机在二维空间维度 (x, y) 上具有高采样能力,即具有高的空间分辨率。四维模型的光场相机的提出,使得视角维度 (θ, φ) 具有较高的分辨率。大量的研究工作结合 (x, y) 和 (θ, φ) 的信息,将空间三维信息投影到了二维焦平面上,获得了较高分辨率的深度信息 (z) ^[40,73]。动态光谱成像系统的提出,使得传感器对离散频率进行积分后得到了较高分辨率的光谱信息 (λ) 。同时,通过每隔一帧在曝光时间内对到达传感器的光通量进行积分 (t) ,也可使时间分辨率达到视频级别的水平^[84]。

然而,与全光函数这一高维连续信号不同,相机是一个低维耦合的离散采样器,即使通过先验信息与后处理计算,也无法完全实现高维高分辨的重构。为了使计算重构突破传统信号的处理极限,实现真实光信号的高维高分辨率成像,需要从本质上突破现有成像模型的局限,将信息优化与计算引入成像过程,对光信号进行优化耦合处理。在实际感知和处理中,作为整合高维信息的光场,会不可避免地存在一定的不确定性,这可理解为系统不确定性、环境不确定性、感知不确定性。具体而言,系统不确定性定义了光通过成像光路、像传感器、光电计算单元、数据传输通道过程中存在的量子测不准现象、像传感器多种噪声来源(散粒噪声和读取噪声)、光电运算引入的编码失真数据传输过程中的丢帧或带宽限制;环境不确定性主要描述被摄场景到成像设备间的不确定性,例如动态场景引入的模糊、复杂场景导致的

不可见、成像介质形成的光散射等;感知不确定性主要存在于对场景的理解过程中,由于感知算法和训练数据的限制,在场景分割、物体检测、异常监测、深度解析等情况下常常发生遗漏或错误等现象。因此,在探索七维光场的感知与处理中,亦需探究光场的“第八维”,即不确定性建模,从而更好地反馈于结构自适应的智能光场成像,提高大场景动态感知的鲁棒性和智能性。

6.2 动态光谱光场成像

相比于传统相机关注空间信息的采集,光场成像中角度信息的加入为光场相机带来了重聚焦、准确三维重建等一系列新的应用。目前,大部分计算光场成像的研究集中于对空间和视角四维信号进行采集和处理上,而全光函数中的光谱及时间维度往往被忽视。在光谱维度上,现有主流方法采用基于红绿蓝三基色成像原理来表示颜色,但是客观世界中的光线蕴含了丰富的光谱信息,光谱成像基于多颜色通道传感器可得到更加完备的多维信息。而时间维度的加入,使得动态光谱光场成像成为可能,并具有广阔的应用前景。

为了实现动态光谱光场成像,在较为成熟的四维光场成像的基础上,需要先探索高分辨率光谱和时间融合的方法。棱镜掩模式光谱视频捕获方法^[85]解决了光谱和时间融合的问题。该方法在不牺牲时间分辨率的前提下,利用棱镜色散原理,提供了高精度光谱,为现有的光谱光场成像方法提供了良好的硬件与理论基础^[85-87]。在此基础上,混合相机式捕获方法和系统的提出,也让动态光谱光场采

集系统具有精准度、实用性及自适应性。基于非结构化光场感知新原理,替换光场相机阵列中子像传感器的感光波段,可以使得光场相机阵列在记录光线高分辨率空间与角度信息的同时,也记录波长和时间信息,因此动态光谱光场成像或将成为现实。动态光谱光场成像在四维光场信息的基础上又新增加了两个维度。现有的动态光场成像主要集中在研究捕获采集工作,如何基于这些多维多尺度信息,探索光场更广泛的应用,值得进一步探索。与此同时,光谱和时间维度的加入,使得光场计算的复杂度极大地增加,如何在保持光场算法精度的同时提升计算效率,也是有待研究的一个方向。

6.3 光电计算光场视觉

现有光场成像系统为采、存、算分离的范式,需要首先进行光电转换并采集存储场景的光场信息,才能通过电子计算执行人工智能算法,然而当前的电子计算面临的摩尔定律增长速度放缓,电子芯片晶体管尺寸逼近物理极限、功耗大,能量效率难以大幅递增的问题,极大地制约了多维多尺度高通量光场视觉信息的感知、重建与处理速度。

Lin 等^[88-89]提出了利用光学衍射属性构造光学神经元以及实现多层神经元三维互联的理论和方法,搭建了全光学衍射深度学习的模型和框架,探索了其在不同机器学习推理任务中的应用,并历史性地实现了光计算在 MNIST 等数据集上的成功分类。在此基础上,下一代光场视觉应以光电计算的理论和方法为指导,开发新一代光电计算智能光场成像设备,实现宽视场高分辨光场信息的高效实时感知、重建与处理。与此同时,将光学人工神经网络与智能光场成像系统紧密结合,将实现全新的人工智能视觉信息处理范式,能够在成像的同时完成多维多尺度光场信息的智能光学计算任务,从而颠覆传统存算分离的速度制约,实现光场采集与计算的无缝衔接,极大地提升光场成像设备对视觉信息的处理速度。

以衍射深度神经网络为出发点,在相干或者部分相干的可见光波段下开展光载视觉信息非线性激活函数的物理实现研究,通过逆向智能设计高集成度的三维微纳结构,获得对光场不同特性(包括时间、空间、角度)的高维调制,提升光学神经元处理信息的能力;研究实现新一代复杂光学人工神经网络的模型、理论与方法,探索三维衍射光学神经网络与光学传感器、光电成像系统的有效集成,构建以光学人工神经网络为基础的新型智能光场成像系统,获

得光场视觉信息的实时感知和重建,实现物体分割、对象识别、场景分类等复杂人工智能任务的超高速推理,并将其应用于自动驾驶、无人机导航、视频监控等领域。此外,还应充分利用光学计算高效率、低功耗的特性,并将其与灵活多变的传统集成电路相结合,以实现速度千倍提升、效率百万倍提升的全新神经形态集成光电计算芯片,为芯片以及当代计算行业带来颠覆性革命,为多维多尺度高通量光场视觉信息的高效处理提供硬件支撑,实现智能光场成像设备的小型化与实时信息处理,从而将极大拓宽光场的应用领域,实现宽视场高分辨光场成像关键技术在大尺度海洋目标搜索、空对地观测以及卫星遥感等领域中的应用。

参 考 文 献

- [1] Gershun A. The light field[J]. *Journal of Mathematics and Physics*, 1939, 18: 51-151.
- [2] Moon P, Spencer D E. The photic field[M]. USA: MIT Press, 1981.
- [3] Adelson H, Bergen J. The plenoptic function and the elements of early vision [M]. USA: MIT Press, 1991: 3-20.
- [4] Levoy M, Hanrahan P. Light field rendering [C] // *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques-SIGGRAPH'96*, August 4-9, 1996, New Orleans, LA, USA. New York: ACM, 1996: 31-42.
- [5] Raytrix. 3D light-field vision [EB/OL]. [2019-11-08]. <https://raytrix.de/>, 2017.
- [6] NGP capital. Pelican imaging [EB/OL]. [2019-11-08]. <https://www.ngpcap.cn/companies/pelican-imaging>.
- [7] Levoy M, Ng R, Adams A, et al. Light field microscopy[J]. *ACM Transactions on Graphics*, 2006, 25(3): 924-934.
- [8] Levoy M. Light fields and computational imaging [J]. *Computer*, 2006, 39(8): 46-55.
- [9] Ihrke I, Restrepo J, Mignard-Debise L. Principles of light field imaging: briefly revisiting 25 years of research [J]. *IEEE Signal Processing Magazine*, 2016, 33(5): 59-69.
- [10] Wetzstein G, Ihrke I, Lanman D, et al. Computational plenoptic imaging [J]. *Computer Graphics Forum*, 2011, 30(8): 2397-2426.
- [11] Chen S E. Quicktime VR: an image-based approach to virtual environment navigation [C] // *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, August 6-11, 1995, Los Angeles, CA, USA. New York: ACM, 1995: 29-38.
- [12] McMillan L, Bishop G. Plenoptic modeling: an

- image-based rendering system[C]//Proceedings of the 22nd annual conference on Computer graphics and interactive techniques, August 6-11, 1995, Los Angeles, CA, USA. New York: ACM, 1995: 39-46.
- [13] Chai J X, Chan S C, Shum H Y, et al. Plenoptic sampling[C]//Proceedings of the 27th annual conference on Computer graphics and interactive techniques-SIGGRAPH'00, July 23-28, 2000, New Orleans, LA, USA. New York: ACM, 2000: 307-318.
- [14] Lin Z C, Shum H Y. A geometric analysis of light field rendering[J]. *International Journal of Computer Vision*, 2004, 58(2): 121-138.
- [15] Buehler C, Bosse M, McMillan L, et al. Unstructured lumigraph rendering[C]//Proceedings of the 28th annual conference on Computer graphics and interactive techniques-SIGGRAPH'01, August 12-17, 2001, Los Angeles, California, USA. New York: ACM, 2001: 425-432.
- [16] Ihrke I, Wetzstein G, Heidrich W. A theory of plenoptic multiplexing[C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE, 2010: 483-490.
- [17] 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.
- [18] Yang J C, Everett M, Buehler C, et al. A real-time distributed light field camera[C]//Proceedings of the 13th Eurographics Workshop on Rendering Techniques, June 26-28, 2002, Pisa, Italy. Frontenex: Eurographics Association, 2002: 77-86.
- [19] 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.
- [20] Wilburn B S, Smulski M, Lee H H K, et al. Light field video camera[J]. *Proceedings of SPIE*, 2001, 4674: 29-36.
- [21] Wilburn B, Joshi N, Vaish V, et al. High-speed videography using a dense camera array[C]//Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. June 27-July 2, 2004, Washington, DC, USA. New York: IEEE, 2004: 8168976.
- [22] Zhang C, Chen T. A self-reconfigurable camera array [C]//Proceedings of the 15th Eurographics Workshop on Rendering Techniques, June 21-23, 2004, Norköping, Sweden. Frontenex: Eurographics Association, 2004: 243-254.
- [23] Chan S C, Ng K T, Gan Z F, et al. The plenoptic video[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2005, 15(12): 1650-1659.
- [24] Liu Y B, Dai Q H, Xu W L. A real time interactive dynamic light field transmission system[C] // 2006 IEEE International Conference on Multimedia and Expo, July 9-12, 2006, Toronto, ON, Canada. New York: IEEE, 2006: 2173-2176.
- [25] ViewPLUS Inc. ProFUSION 25 5×5 camera array system [EB/OL]. [2019-11-08]. <http://www.viewplus.co.jp/products/profusion25/index-%20e.html,%202009>.
- [26] The (New) Stanford light field archive [EB/OL]. [2019-11-08]. <http://lightfield.stanford.edu/lfs.html>.
- [27] Unger J, Wenger A, Hawkins T, et al. Capturing and rendering with incident light fields[C]//Proceedings of the 14th Eurographics workshop on Rendering Techniques, June 25-27, 2003, Leuven, Belgium. Frontenex: Eurographics Association, 2003: 141-149.
- [28] 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): 73.
- [29] Dansereau D G, Schuster G, Ford J, et al. A wide-field-of-view monocentric light field camera[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 3757-3766.
- [30] Liang C K, Lin T H, Wong B Y, et al. Programmable aperture photography: multiplexed light field acquisition[J]. *ACM Transactions on Graphics*, 2008, 27(3): 55.
- [31] Ihrke I, Stich T, Gottschlich H, et al. Fast incident light field acquisition and rendering [J]. *Journal of WSCG*, 2008, 16(1/2/3): 25-32.
- [32] Taguchi Y, Agrawal A, Ramalingam S, et al. Axial light field for curved mirrors: reflect your perspective, widen your view[C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE, 2010: 499-506.
- [33] Lippmann G. La photographie integrale[J]. *Academie des Sciences*, 1908, 146: 446-451.
- [34] 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.
- [35] Ng R, Levoy M, Brédif M, et al. Light field

- photography with a hand-held plenoptic camera[R]. USA: Computer Science Technical Report CSTR, 2005.
- [36] LightField Forum. Lytro [EB/OL]. [2019-11-08]. <http://lightfield-forum.com/lytro/>.
- [37] Wei L Y, Liang C K, Myhre G, et al. Improving light field camera sample design with irregularity and aberration[J]. *ACM Transactions on Graphics*, 2015, 34(4): 152.
- [38] Georgiev T, Lumsdaine A. Superresolution with plenoptic camera 2.0[R]. USA: Adobe Systems Incorporated, 2009.
- [39] Horstmeyer R, Euliss G, Athale R. Flexible multimodal camera using a light field architecture[C] // 2009 IEEE International Conference on Computational Photography (ICCP), April 16-17, 2009, San Francisco, CA, USA. New York: IEEE, 2009: 11499063.
- [40] Levin A, Fergus R, Durand F, et al. Image and depth from a conventional camera with a coded aperture[J]. *ACM Transactions on Graphics*, 2007, 26(3): 70.
- [41] Manakov A, Restrepo J F, Klehm O, et al. A reconfigurable camera add-on for high dynamic range, multispectral, polarization, and light-field imaging [J]. *ACM Transactions on Graphics*, 2013, 32(4): 47.
- [42] Lanman D, Crispell D, Wachs M, et al. Spherical catadioptric arrays: construction, multi-view geometry, and calibration [C] // Third International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06), June 14-16, 2006, Chapel Hill, NC, USA. New York: IEEE, 2006: 10285053.
- [43] Levoy M, Zhang Z, McDowall I. Recording and controlling the 4D light field in a microscope using microlens arrays[J]. *Journal of Microscopy*, 2009, 235(2): 144-162.
- [44] Prevedel R, Yoon Y G, Hoffmann M, et al. Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy[J]. *Nature Methods*, 2014, 11(7): 727-730.
- [45] Nöbauer T, Skocek O, Pernía-Andrade A J, et al. Video rate volumetric Ca^{2+} imaging across cortex using seeded iterative demixing (SID) microscopy [J]. *Nature Methods*, 2017, 14(8): 811-818.
- [46] 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.
- [47] Mignard-Debise L, Ihrke I. Light-field microscopy with a consumer light-field camera[C] // 2015 International Conference on 3D Vision, October 19-22, 2015, Lyon, France. New York: IEEE, 2015: 335-343.
- [48] Georgiev T G, Zheng K C, Curless B, et al. Spatio-angular resolution tradeoffs in integral photography [C] // Proceedings of the Eurographics Symposium on Rendering Techniques, Nicosia, Cyprus. Frontenex: Eurographics Association, 2006: 263-272.
- [49] Veeraraghavan A, Raskar R, Agrawal A, et al. Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing [C] // SIGGRAPH'07 ACM SIGGRAPH 2007 papers, August 5-9, 2007, San Diego, California. New York: ACM, 2007: 69.
- [50] Lanman D, Raskar R, Agrawal A, et al. Shield fields: modeling and capturing 3D occluders[J]. *ACM Transactions on Graphics*, 2008, 27(5): 131.
- [51] Amit A, Mark A. Compressive light field imaging [J]. *Proceedings of SPIE*, 2010, 7690: 76900Q.
- [52] Antipa N, Necula S, Ng R, et al. Single-shot diffuser-encoded light field imaging [C] // 2016 IEEE International Conference on Computational Photography (ICCP), May 13-15, 2016, Evanston, IL, USA. New York: IEEE, 2016.
- [53] Pégard N C, Liu H Y, Antipa N, et al. Compressive light-field microscopy for 3D neural activity recording [J]. *Optica*, 2016, 3(5): 517-524.
- [54] Cohen N, Yang S, Andalman A, et al. Enhancing the performance of the light field microscope using wavefront coding[J]. *Optics Express*, 2014, 22(20): 24817-24839.
- [55] 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.
- [56] Wang Y P, Wang L C, Kong D H, et al. High-resolution light field capture with coded aperture[J]. *IEEE Transactions on Image Processing*, 2015, 24(12): 5609-5618.
- [57] Kopf J, Uyttendaele M, Deussen O, et al. Capturing and viewing gigapixel images[J]. *ACM Transactions on Graphics*, 2007, 26(3): 93.
- [58] Gigapan [EB/OL]. [2019-11-08]. <http://www.gigapan.com/>.
- [59] Pirk S, Cohen M F, Deussen O, et al. Video enhanced gigapixel panoramas [C] // SIGGRAPH Asia 2012 Technical Briefs on-SA'12, November 28-December 1, 2012, Singapore, Singapore. New York: ACM, 2012: 7.
- [60] Cossairt O S, Miao D, Nayar S K. Gigapixel computational imaging [C] // 2011 IEEE International

- Conference on Computational Photography (ICCP), April 8-10, 2011, Pittsburgh, PA, USA. New York: IEEE, 2011: 11962821.
- [61] Cossairt O S, Miao D, Nayar S K. Scaling law for computational imaging using spherical optics [J]. *Journal of the Optical Society of America A*, 2011, 28(12): 2540-2553.
- [62] Brady D J, Gehm M E, Stack R A, et al. Multiscale gigapixel photography[J]. *Nature*, 2012, 486(7403): 386-389.
- [63] Kittle D S, Daniel L, David J. Automated calibration and optical testing of the AWARE-2 gigapixel multiscale camera [J]. *Proceedings of SPIE*, 2013, 8660: 866006.
- [64] Kittle D S, Marks D L, Son H S, et al. A testbed for wide-field, high-resolution, gigapixel-class cameras[J]. *Review of Scientific Instruments*, 2013, 84(5): 053107.
- [65] Pang W, Brady D J. Parallel MMS gigapixel imagers [C] // *Computational Optical Sensing and Imaging*, June 26-29, 2017, San Francisco, California, United States. Washington, D. C.: OSA, 2017: CM2B.3.
- [66] Schuster G M, Agurok I P, Ford J E, et al. Panoramic monocentric light field camera[C] // *International Optical Design Conference*, July 9-13, 2017, Denver, Colorado, United States. Washington, D. C.: OSA, 2017: ITh4A.5.
- [67] Wang Y W, Liu Y B, Heidrich W, et al. The light field attachment: turning a DSLR into a light field camera using a low budget camera ring [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2017, 23(10): 2357-2364.
- [68] Yuan X Y, Fang L, Dai Q H, et al. Multiscale gigapixel video: a cross resolution image matching and warping approach [C] // *2017 IEEE International Conference on Computational Photography (ICCP)*, May 12-14, 2017, Stanford, CA, USA. New York: IEEE, 2017: 16969274.
- [69] 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.
- [70] Zhang Z, Liu Y B, Dai Q H. Light field from micro-baseline image pair [C] // *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 3800-3809.
- [71] Pearson J, Brookes M, Dragotti P L. Plenoptic layer-based modeling for image based rendering [J]. *IEEE Transactions on Image Processing*, 2013, 22(9): 3405-3419.
- [72] Penner E, Zhang L. Soft 3D reconstruction for view synthesis [J]. *ACM Transactions on Graphics*, 2017, 36(6): 235.
- [73] Chaurasia G, Duchene S, Sorkine-Hornung O, et al. Depth synthesis and local warps for plausible image-based navigation [J]. *ACM Transactions on Graphics*, 2013, 32(3): 30.
- [74] Kalantari N K, Wang T C, Ramamoorthi R. Learning-based view synthesis for light field cameras [J]. *ACM Transactions on Graphics*, 2016, 35(6): 193.
- [75] Zhou T, Tucker R, Flynn J, et al. Stereo magnification: learning view synthesis using multiplane images [J/OL]. (2018-05-24) [2019-11-08]. <https://arxiv.org/abs/1805.09817>.
- [76] Mildenhall B, Srinivasan P P, Ortiz-Cayon R, et al. Local light field fusion: practical view synthesis with prescriptive sampling guidelines [J/OL]. (2019-05-02) [2019-11-08]. <https://arxiv.org/abs/1905.00889>.
- [77] Chai J X, Ton X, Chan S C, et al. Plenoptic sampling [C] // *Proceedings of the 27th annual conference on Computer graphics and interactive techniques-SIGGRAPH'00*, July 23-28, 2000, New Orleans, LA, USA. New York: ACM, 2000: 307-318.
- [78] Shi L X, Hassanieh H, Davis A, et al. Light field reconstruction using sparsity in the continuous Fourier domain [J]. *ACM Transactions on Graphics*, 2014, 34(1): 12.
- [79] Vagharshakyan S, Bregovic R, Gotchev A. Light field reconstruction using shearlet transform [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(1): 133-147.
- [80] Vagharshakyan S, Bregovic R, Gotchev A. Accelerated shearlet-domain light field reconstruction [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2017, 11(7): 1082-1091.
- [81] Yoon Y, Jeon H G, Yoo D, et al. Learning a deep convolutional network for light-field image super-resolution [C] // *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 57-65.
- [82] Yeung H W F, Hou J H, Chen J, et al. Fast light field reconstruction with deep coarse-to-fine modeling of spatial-angular clues [M] // Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018*. Lecture notes in computer science. Cham: Springer, 2018, 11210: 138-154.
- [83] Wu G C, Liu Y B, Fang L, et al. Light field reconstruction using convolutional network on EPI and extended applications [J]. *IEEE Transactions on*

- Pattern Analysis and Machine Intelligence, 2019, 41(7): 1681-1694.
- [84] Dai Q H, Suo J L, Ji X Y, et al. Computational photography: computational capture of plenoptic visual information[M]. Beijing: Tsinghua University Press, 2016: 11.
戴琼海, 索津莉, 季向阳, 等. 计算影像学: 全光视觉信息的计算采集[M]. 北京: 清华大学出版社, 2016: 11.
- [85] Zhu K, Xue Y J, Fu Q, et al. Hyperspectral light field stereo matching[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(5): 1131-1143.
- [86] Holloway J, Mitra K, Koppal S J, et al. Generalized assorted camera arrays: robust cross-channel registration and applications[J]. IEEE Transactions on Image Processing, 2015, 24(3): 823-835.
- [87] Xiong Z W, Wang L Z, Li H Q, et al. Snapshot hyperspectral light field imaging[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 6873-6881.
- [88] Lin X, Rivenson Y, Yardimci N T, et al. All-optical machine learning using diffractive deep neural networks[J]. Science, 2018, 361(6406): 1004-1008.
- [89] Yan T, Wu J M, Zhou T K, et al. Fourier-space diffractive deep neural network[J]. Physical Review Letters, 2019, 123(2): 023901.