

压缩光场成像: 原理、方法及应用

李明翰^{1,2}, 魏凯^{1,2}, 封晓华^{1,2}

(1. 浙江大学 光电科学与工程学院, 杭州 310027; 2. 浙江大学 杭州国际科创中心, 杭州 311215)

摘要: 光场成像能够同时记录三维场景中光线的空间与角度信息, 从而实现传统二维成像难以具备的视点变化与深度感知等功能。随着计算摄影与成像技术的不断进步, 光场相机已从概念验证阶段发展到多种采集方案并存的实用化阶段。本文系统回顾了光场成像的发展历程与基本理论, 阐述全光函数及光场的四维参数化模型; 总结了现有光场数据获取的主要实现途径及其代表性研究, 并着重介绍了压缩光场成像的基本原理与研究现状。最后, 文章总结了光场成像在各领域的典型应用, 如 3D 成像、增强现实/混合现实(AR/MR)以及工业检测等。然而, 尽管取得了诸多进展, 光场成像仍面临着数据存储、实时处理、硬件复杂性等方面的挑战。展望未来, 随着光学编码技术的不断创新和深度学习算法的优化, 压缩光场成像有望在多个领域实现更加广泛的应用。

关键词: 光场成像; 压缩光场成像; 相机阵列; 微透镜阵列; 深度学习

中图分类号: TN 278 **文献标志码:** A

Compressive light field imaging: principles, methods, and applications

LI Minghan^{1,2}, WEI Kai^{1,2}, FENG Xiaohua^{1,2}

(1. College of Optical Science and Engineering, Zhejiang University, Hangzhou 310027, China; 2. ZJU-Hangzhou Global Scientific and Technological Innovation Center, Zhejiang University, Hangzhou 311215, China)

Abstract: Light field imaging records both spatial and angular information of light rays in a 3D scene, enabling capabilities such as viewpoint variation and depth perception that are not possible with traditional 2D imaging. As computational photography and imaging technologies advance, light field cameras have evolved from conceptual designs to practical applications through various acquisition techniques. This review systematically examined the development history and fundamental concepts of light field imaging, including the plenoptic function and four-dimensional parameterization of light fields. We summarized different implementation approaches for light field data acquisition and their

收稿日期: 2025-08-06

基金项目: 国家自然科学基金青年基金资助项目(62205302)

第一作者: 李明翰(1999—), 男, 博士研究生. 研究方向: 计算成像、光场摄影. E-mail: minghanli@zju.edu.cn

通信作者: 封晓华(1988—), 男, 研究员. 研究方向: 计算成像、光场成像、三维视觉. E-mail: xiaohuafeng@zju.edu.cn

引文格式: 李明翰, 魏凯, 封晓华. 压缩光场成像: 原理、方法及应用[J]. 上海理工大学学报, 2025, 47(6): 622-639.

Citation: LI Minghan, WEI Kai, FENG Xiaohua. Compressive light field imaging: principles, methods, and applications[J]. Journal of University of Shanghai for Science and Technology, 2025, 47(6): 622-639.

representative research, with particular emphasis on the basic principles and current state of compressive light field imaging. Finally, the article summarized the typical applications of light field imaging in various fields, such as 3D imaging, augmented reality/mixed reality (AR/MR), and industrial inspection. However, despite numerous advancements, light field imaging still faces challenges in data storage, real-time processing, and hardware complexity. Looking ahead, with continuous innovation in optical encoding technology and optimization of deep learning algorithms, compressive light field imaging is expected to achieve broader applications in multiple fields.

Keywords: *light field imaging; compressive light field imaging; camera array; micro-lens array; deep learning*

1 光场成像的发展历程与基础概念

光场成像技术被认为是传统应用光学的一个革命性的技术,是当下最有前景的三维成像技术之一。光场成像通过记录光线的空间位置与传播方向,保留了更多维度的光子信息,使得数字重聚焦^[1-3]、景深调节^[4-6]、三维立体观察^[7-8]、精确深度估计^[9-12]以及自由视点漫游^[13-14]等后期图像处理功能成为可能。

光场(light field)最早由 Lippman 等^[15]于 1908 年提出,之后 Gershun^[16]在 1936 年利用光场描述了空间中光线的辐射分布。1991 年,Adelson 等^[17]提出了全光函数(plenoptic function)的概念,用一个 7 维函数 $P(x,y,z,\theta,\varphi,\lambda,t)$ 描述场景中任意点的光

线随着位置、方向、波长和时间的变化。在静止场景和恒定光照假设下,全光函数可简化为 5 维(去除时间和波长)^[18]。进一步假定光线在自由空间中的强度沿直线传播不变,则可用 4 维函数即“光场”来表示。如图 1 所示,典型的四维光场采用 Levoy 等^[3]提出的“两平面参数化”光场模型:在空间中引入两个平行平面,一个作为摄像机主平面(s,t),一个作为图像平面(u,v),那么一条光线可由其在两个平面上的交点坐标(u,v,s,t)唯一确定。这一定义将光场表示为在空间平面上从不同视角看到的众多视图图像集合。光场也可视为一组具有空间和角度维度的影像数据,其特殊切片——EPI(epipolar plane image)提供了分析场景几何的有力工具^[19]。

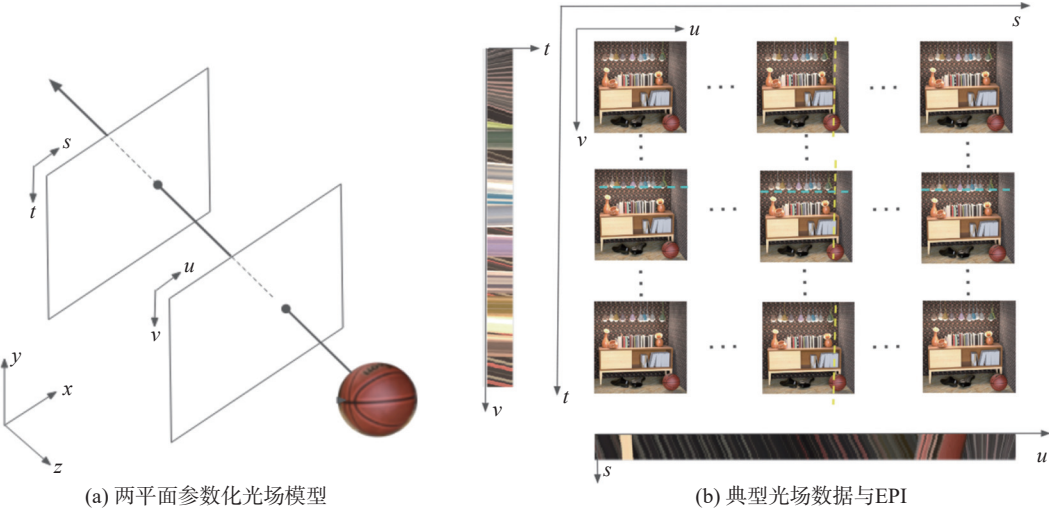


图 1 光场示意图^[3]
Fig.1 Light field schematic diagram

光场成像的主要目的是突破传统 2D 成像中的固有局限性。传统相机通过透镜将三维场景映射为二维图像,无法保留光线的入射角度信息,因

而失去了视差信息与三维观测的能力^[20-22]。为记录光线角度信息,研究人员提出了众多的方案。Lippmann^[15]的积分摄影通过在感光板前放置透镜

阵列以记录不同方向来的光线。Adelson 等^[23]于 1992 年通过在相机主镜头与感光传感器之间加入微透镜阵列,将来自不同方向的光线映射至不同像素,从而在一次曝光中采集到四维光场数据。进入 21 世纪后,数字传感器的发展极大推动了光场相机的实用化:Stanford 大学的 Ng 等^[24]在 2005 年研制出了手持式光场相机,并提出了数字重聚焦等应用。Lytro 公司在 2011 年推出了消费级光场相机;Raytrix 公司则推出了多焦距微透镜光场相机,用于工业和科学成像。经过 20 余年的快速演进,光场相机硬件^[25-26]、信号处理算法^[27],以及编码传输已形成“采集-压缩-重建-显示”的完整技术栈^[28]。其高维度数据所带来的灵活图像后处理与三维感知能力,使光场成像成为计算摄影与计算成像领域最活跃、最成熟的研究方向之一^[29]。

然而,光场成像技术仍面临着空间与角度分辨率的折衷这一核心挑战。由于传感器像素数有限,直接记录四维光场意味着必须在空间分辨率(图像清晰度)和角度分辨率(视差细节)之间作出权衡:光场相机通常通过牺牲空间分辨率来换取角度信息。例如,基于微透镜阵列的光场相机将传感器像素划分成微透镜下的子图像,从而降低了每幅视图的像素数量。如何在获取高维光场信息的同时保持较高的成像分辨率、信噪比和实时性能,是光场成像技术未来发展的主线。本文首先介绍传统光场成像中不同的光场获取方法及其特点,继而介绍压缩光场成像在突破空间-角度分辨折衷以及降低采样数据方面的独特优势^[30-32]。

2 传统光场成像的不同实现方式

光场数据的获取可以通过多种技术途径来实现,主要区别在于空间和角度采样的策略不同。本节将介绍几种典型方案(单相机机械扫描、相机阵列以及微透镜阵列),并比较它们的原理和特点。此方案在采集 4D 光场时,往往在时间、空间、角度等不同维度之间作出了权衡,通过牺牲某个维度上的性能来提升另一个维度上的性能。

2.1 单相机逐视点移动采集

采集四维光场最直接的办法是使用单个相机通过机械扫描在不同视点拍摄多张照片,再将其组合得到光场。这种方法实质是在时间上对光场进行采样:相机可以沿直线轨道平移,或者在二

维平面上扫描拍摄,每个位置获得场景的一个视角。最终将多视角图像排列,即可形成光场的 (s,t) 维采样。该方式无需特殊光学器件,使用常规相机即可获取高分辨率的多视图图像,典型案例如 Stanford 大学在 1996 年的工作^[3]。其主要优点是空间分辨率不受影响,每张照片都是全传感器分辨率,并且灵活性高,可以根据需要采集任意视点分布的数据。然而,缺点也很明显:首先,逐视点拍摄耗时较长,不适用于动态场景;其次,多张图像的对准和校准较复杂,可能需要借助结构光或标定板来精确对齐视差。因此,单相机扫描通常用于静态场景的光场采集或实验室条件下验证。

近年来也有研究将手持相机的随机移动与 SLAM(同步定位与地图构建)技术结合,实现对散乱视点图像的融合以生成光场^[33]。总体而言,单相机移动方案设备简单但采集效率低,在实用中多被其他方案所取代。

2.2 多相机阵列采集

利用多相机阵列可以同步获取场景的多视角图像,实现光场采集的并行化。典型做法是将若干摄像机按规则排列(如平面阵列或圆形阵列),同时拍摄同一场景,各相机相当于采样光场不同的角度位置。图 2 所示为通过相机阵列实现光场采集的典型场景。多相机阵列能够在瞬间获得完整的光场,对于动态场景和视频光场采集非常有效^[34-36]。其优势在于:由于每个相机独立成像,可充分利用每台相机的传感器分辨率,光场的空间分辨率由单台相机决定;同时,角度分辨率由相机数量决定,可以较灵活地权衡覆盖视场和视角密度。此外,可以通过增加相机数量实现大光圈合成效果。例如,同时开启阵列中所有相机再叠加,可模拟一个等效大孔径相机进行光学合成孔径成像,实现景深外的物体虚化或透过遮挡观察^[37-38]。

多相机阵列的主要劣势是成本高、系统复杂,需要大量相机、同步触发和数据传输,以及统一色彩和几何校准等条件。阵列尺寸受限于物理安装,会影响可采集基线范围。同时,阵列通常固定安装,缺乏便携性。Stanford 计算成像阵列是著名的实例^[39]:由 100 台摄像机组成 10×10 平面阵列,实现高分辨率、高帧率的光场视频采集。他们利用该阵列实现了每秒 30 帧的光场视频,以及通过视差插值实现高速摄影和慢动作效果等应

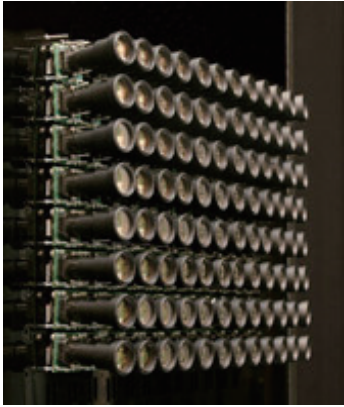


图 2 相机阵列实现光场采集^[39]

Fig.2 Light field capture using a camera array

用。多相机阵列适合专业应用(如电影特效中的“子弹时间”镜头)，在科研中也常用于获取光场基准数据。但由于造价和体积问题，消费级产品中较少直接采用大规模阵列，而是通过其他方式模拟多视角采样。

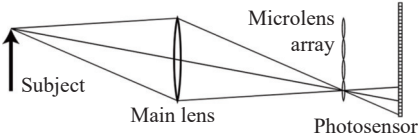
2.3 微透镜阵列光场相机

微透镜阵列光场相机(简称微透镜光场相机或平面光场相机)是在单相机内部实现光场采集的经典方案。其结构特点是在主镜头与成像传感器之间引入一块由成百上千个微小透镜组成的微透镜阵列(MLA)，每个微透镜充当一个小型“子相机”。主镜头将场景成像到微透镜阵列平面时，每个微透镜会将通过该透镜的光线进一步映射到传感器上，形成一个微小的子图像。这些子图像又称微透镜图像阵列，通常每个微透镜对应大约10~20个像素。一个微透镜阵列相机获得的原始图像中，肉眼可见密集排列的微小影像，每个微透镜下不同像素对应不同的光线传播方向。通过对微透镜图像进行解码，可以将每个微影像的位置映射为光场中的一个视角 (s,t) 。影像内的像素位置映射为该视角下的坐标 (u,v) ，从而得到一系列子孔径视图(即不同视角的图像)^[40-41]。

微透镜阵列光场相机的原理是以空间分辨率换取角度分辨率：传感器上的像素同时记录光线传播的空间和角度信息，使每张子视图的空间分辨率大大降低。其优点是单次曝光即可获取四维光场，非常适合动态场景和手持拍摄。该方案自Adelson等最初提出后，Ng等^[24]继续完善了设计(取消了中继透镜等)，有效增加了像素利用率，大幅提高了光场相机的实用性。2000年后的Lytro和Raytrix光场相机都属于此类设计。如图3



(a) Lytro生产的商用光场相机实物图



(b) 微透镜阵列光场相机光路原理图

图 3 微透镜阵列光场相机^[24]

Fig.3 Micro-lens array light field camera

所示，Lytro 光场相机只有一个手掌大小(图3(a))，其原理(图3(b))是利用主镜头将物体成像到微透镜平面，每个微透镜再将来自不同方向的光线分开聚焦到图像传感器上，从而记录角度信息。Lytro 相机搭载了1100万像素传感器，但最终生成的每张子视图不到10万像素，导致单视图清晰度较差。事实上，基于微透镜的光场成像方案长期面临着图像分辨率难以有效提升的难题。

2022年，清华大学戴琼海院士课题组研发了Meta sensor，其基本模型也是基于微透镜阵列的光场相机，但是采用了创新性的扫描光场成像的办法，起到增加光场分辨率的作用^[42]。如图4所示：扫描光场成像通过在微透镜阵列前引入像素级空间扫描，增大空间采样率；同时，通过控制微透镜尺寸引入额外的光学相干性，使得高分辨率合成孔径成为可能，最终实现高分辨光场成像。为了减少传统微透镜阵列光场相机光场重建时的伪影，扫描光场成像还在每个微透镜周边涂上铬薄膜，以阻挡穿过每个微透镜方格周围角落的光。这种办法减少了光学传递函数中的零点，促进无伪影重建，这对于通用成像场景至关重要。

总体而言，3种传统光场采集方式在分辨率、采集效率及系统复杂度方面各有优势。单相机扫描方式空间分辨率很高，但是大大牺牲了时间分辨率，不适用于动态场景；多相机阵列的光场成像可以进行实时采集，时间、空间、角度分辨率理论上都可以很高，但系统复杂、体积庞大、成本较高等因素限制了其实际的应用；微透镜阵列相机结构紧凑、适合动态场景，其作为目前最成

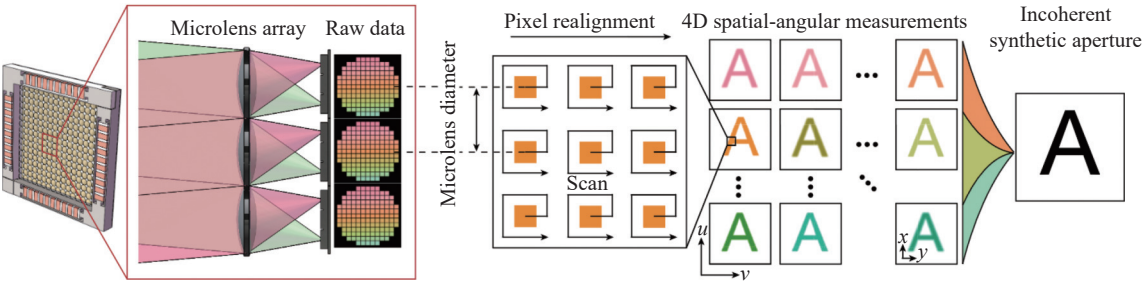


图 4 扫描光场成像^[42]

Fig.4 Scanning light field imaging

熟的光场相机方案，已经在科研和工业领域得到广泛应用。很多改进型光场成像也都基于该方案。比如，Raytrix^[25]使用了不同焦距的微透镜的光场相机，可以兼顾成像景深和空间分辨率。但是受限于空间-角度折衷问题，微透镜阵列光场相机难以记录高角度、空间分辨率的光场。上述局限性促使研究者们发展新型光场成像技术，突破传统硬件在分辨率与采样效率上的瓶颈。

3 压缩光场成像：成像技术与重建算法的结合

光场的四维数据给传统相机带来了空间分辨率与角度分辨率的折衷：传感器像素有限，直接记录四维光场需在图像清晰度和视差细节间分配资源，若使用扫描或多相机方案又会很大程度上增加系统复杂性。压缩光场通过对传感器像素的高效复用，可以突破空间分辨率与角分辨率权衡的问题^[32]。压缩光场成像的核心思想是：以光学编码实现对高维光场的稀疏投影，以算法重建恢复完整光场^[43-44]。其基本原理是通过特殊光学元件将四维光场压缩编码成二维图像^[45-46]，再借助稀疏性先验^[31,47]或深度学习模型^[48]从这张二维测量中重建光场。图像传感器记录的是混合了不同视角光线的信息，后端算法根据预先假设的稀疏表示或训练的神经网络，将其解码出原始的光场数据^[49]。这样的成像流程突破了微透镜阵列相机“空间换角度”的固定折衷，通过“以算代采”实现少量甚至单张拍摄获取高分辨率光场，是下一代光场成像技术的发展方向。

3.1 基于稀疏重建的压缩光场成像

自 2006 年以来，压缩感知理论以“少量测量 + 稀疏先验 + 高效优化”思路重塑了信号获取与重建范式，推动了医学成像、无线通信、计算成

像、机器学习等领域的采样优化与算法创新^[50-57]。早期的压缩光场成像深受稀疏重建算法发展的影响，通过在光路中添加调制元件对光场进行编码，再利用光场的稀疏性对测量结果进行反解得到光场^[58-59]。2007 年，Veeraraghavan 等^[44]学者提出了混频光场理论框架，通过传统相机光学路径中引入可编程衰减掩模，实现对四维光场信息的可逆调制与高效编码。与传统光场相机不同，该工作创新性地摒弃了通过弯曲光线来捕获角度信息的思路，转而采用掩模对光线进行选择性的衰减，将高维光场信息压缩到二维传感器上。其重建算法简单，只需将传感器上的二维傅里叶变换的图块重新排列成四维平面并计算逆傅里叶变换即可。

2010 年，Ashok 等^[59]创新性地设计了两种压缩光场成像架构，仅需采集相对较少的测量数据即可实现高分辨率光场重建，显著提高了成像系统的效率。如图 5 所示，他们将压缩光场分为角度压缩光场与空间压缩光场。两种方案的核心思想都是用一片可编程振幅掩模对 4D 光场进行有控制的线性混叠，将高维信号投影为少量 2D 测量。其中，角度压缩光场成像将掩膜置于孔径平面，对进入成像系统的多束角向光线进行加权编码。而空间压缩光场成像将掩膜放置在图像平面前后，对同一视角的空间像素进行编码，使每次测量同时混合多个空间采样点，再通过稀疏先验与优化算法重建完整光场。他们的工作为压缩光场成像技术奠定了坚实的理论基础与实践经验，启发了后续大量的压缩光场重建方法。

在角度压缩光场成像技术中，最具代表性的工作是 Marwah 等^[60]在 2013 年提出的一种压缩光场相机架构。他们设计了优化的掩膜来进行光场的压缩采样，并基于过完备字典学习将光场稀疏表示为一组“光场原子”，再通过优化算法从单

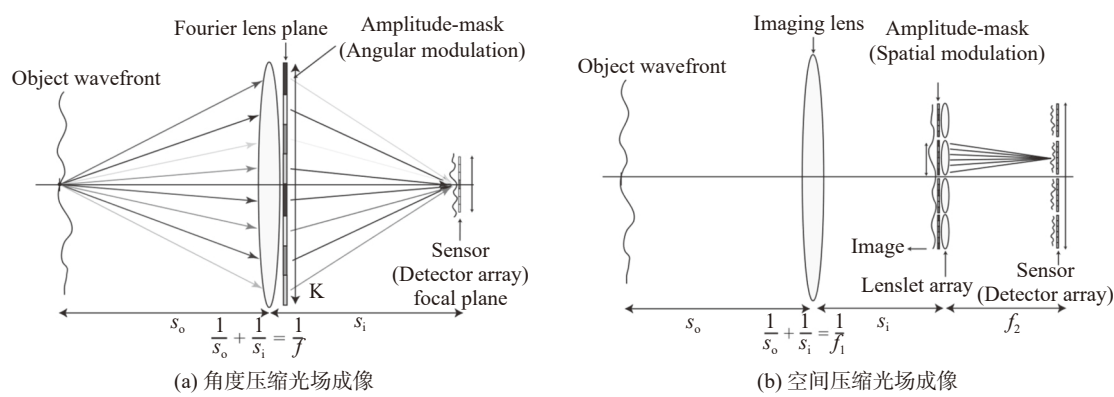


图 5 两种典型压缩光场成像光路^[59]

Fig.5 Two typical imaging optical paths of compressive light field imaging

幅编码图像重建出高分辨率光场。图 6 (a) 为 Marwah 等的实验装置，图 6(b) 展示了原始采集的光场投影数据以及算法解码后的图片。图中： s_0 为物距； s_i 为像距； f 为透镜焦距； f_1 为主透镜焦距； f_2 为微透镜焦距。其优势在于：只需一次曝光，就能够恢复出相当于多视图阵列或高像素微透镜相机才能获得的光场细节。特别是，它使得光场采集的空间分辨率接近传感器原生分辨率，显著提高了光场数据的清晰度。

然而，Marwah 等工作的不足在于：他们必须显式地学习一组完备稀疏字典，这个过程不仅复杂，而且泛化性较差，由于数据集的限制问题，难以泛化到所有光场场景的重建问题中。此外，根据稀疏重建理论，利用信号的稀疏性来重建光场所需的采样次数有理论下限，可以表示为 $m \sim k \log(d/k)$ 。式中： m 表示采样次数； k 表示信号的稀疏度； d 表示信号向量的维度。如果重建的光场对象角、空间分辨率都更大，采用稀疏性先验进行单曝光压缩重构的方法会迅速失效；同时，对于噪声较大的情况，重建质量也会下降。尽管如此，Marwah 等工作有效推动了压缩光场

成像技术的实用化进展。

除了字典稀疏表示，一些研究还探索了光场的低秩性质和稀疏分量结合，以张量分解等方式来提高重建质量和效率^[61-65]。总之，基于稀疏重建的方法代表了人们早期对于压缩光场重建的探索过程，其思想简单，在特定场景下效果也较好。但是，基于稀疏重建的方法一般通过迭代优化实现，高精度重建往往需要较长计算时间(通常每张光场图像需数十分钟甚至数小时)。因此，这类方法在实现高分辨率的同时，如何加速重建以适应视频帧率成为一大挑战。

3.2 基于深度学习的压缩光场成像

近十年来，深度学习在算力提升 (GPU/TPU)^[66-67]、大规模数据集、模型结构(从卷积神经网络^[68-70]到 Transformer^[71]、自监督学习^[72-73]、扩散模型^[74-79])等多维度创新的共同推动下，从特征工程转向端到端学习，使视觉、语言、语音、多模态理解^[80-82]等任务的性能大幅跃升。深度学习的发展也为压缩光场重建提供了新思路。

CNN 等神经网络能够从大量训练数据中学到光场重建的映射关系，大幅提升重建速度和效

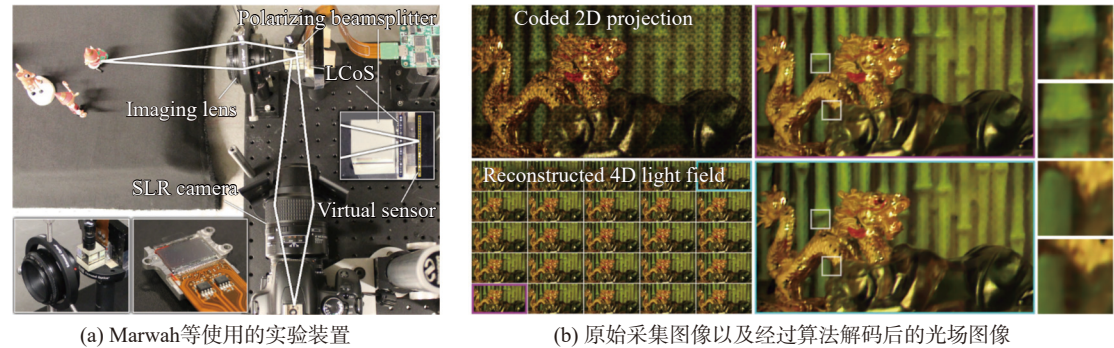


图 6 掩膜编码光场成像原理示意图^[60]

Fig.6 Schematic of mask-coded light field imaging

果。例如，Gupta 等^[48](2017)构建了包含自编码器和 4D 卷积分支的双路网络，从单张编码图像直接推断出高分辨率光场。如图 7 所示，该方法使用神经网络的直接推理替代了先前利用稀疏字典通过多次迭代来重建光场的办法。作者从 4 个合成光场数据集中提取了 50000 个光场块，从 Lytro Illuim 相机捕获的 100 个真实光场数据集中提取了 500000 个光场块(所有光场块的大小都为 $9\times9\times5\times5$)，正向建模光场捕获过程以模拟编码捕获，并以 85 : 15 的比例分割训练集和验证集。在 NVIDIA Titan X GPU 上训练了 60 轮大约 7 h，得到了一个端到端光场重建神经网络。该网络可以直接从编码图中重建出 4D 光场，相比于迭代办法，重建时间从约 35 min 缩短至 6~7 min。

相较于以局部卷积为核心的 CNN，Transformer 通过自注意力实现全局依赖建模、天然的并行计算与可扩展性：在大规模数据/预训练或更高效的

训练策略下，基于 Transformer 的图像处理模型在各类图像处理任务中先后达到或超过了最好的 CNN 模型^[83-84]。受益于 Transformer 的发展，Hua 等提出利用 Transformer 架构同时建模光场的空间和角度依赖关系，设计专门的角度注意力模块来重建稠密视图这类方法。由于善于捕捉全局相关性，如图 8(a) 所示，所提出的 LFRTR 包括 3 个模块：初始特征提取、空间-角度特征合并和视图重建。图 8(b) 的实验结果表明这种办法在合成新视角和提升视差细节方面表现出色^[85]，相较于其他方法能够获得更小的重构误差。

Gao 等^[86]最近探索了基于扩散模型的光场合成，他们提出了 Lfdiff 方法，仅采用单个 RGB 图像作为输入，利用预训练的单眼深度估计网络的视差，将视差信息显式纳入条件信号中。如图 9(a) 所示，迭代 T 时间步长后，随机高斯噪声被去噪为高质量的光场。其网络包括两个部分：

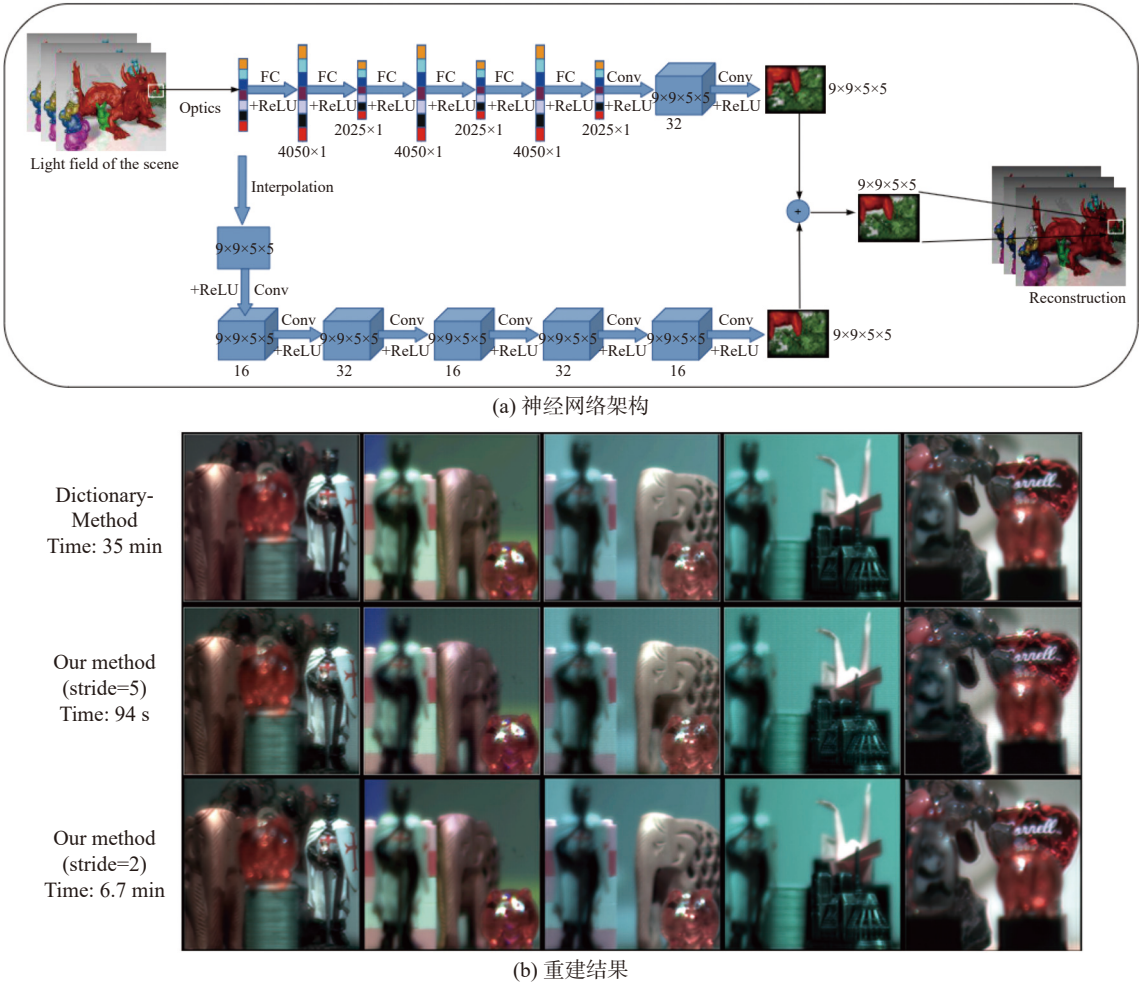


图 7 使用深度学习进行光场重建的网络架构与重建结果^[83]

Fig.7 Network architecture and reconstruction results of light field reconstruction based on deep learning

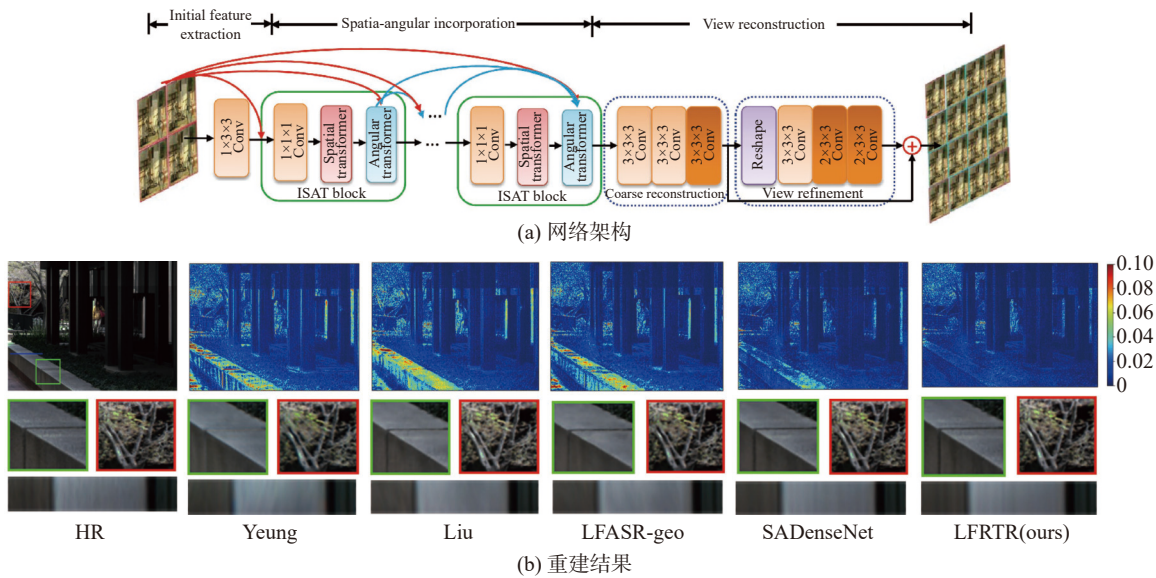


图 8 基于 Transformer 架构的压缩光场重构^[85]
Fig.8 Compressive light field reconstruction based on Transformer

position-aware warping condition scheme 与 disentangled noise estimation network。前者是为了将光场的每个子孔径视图都限制在深度估计网络得到的视差条件中，后者是解耦式的噪声预测网络，将光场的空间与角度特征分离建模，增强多视角一致性的表达能力。图 9(b) 展示了最后的效果，与其他方法相比有着更清晰的细节。

总体而言，深度学习方法通过数据驱动的先验有效提升了压缩光场成像的重建速度和鲁棒性，相比于稀疏重建的迭代重建方法，神经网络的端到端推理能力使实时光场视频采集成为可能。不过，这类方法往往依赖大规模训练数据，如何保证对未知场景的泛化能力仍是需要关注的问题。

3.3 新型光场编码采集与重构

压缩光场成像中的两大模块：光场编码采集^[45]与图像重构，通常都是单独设计与优化的。而随着可微分光学建模与图像重构算法的发展，协同优化光学系统与图像重构算法被证明可以获得更高质量的光学成像与感知效果。2020 年，Guo 等^[46]提出将可学习掩膜的光学编码与深度重建网络端到端联合优化。如图 10(a) 所示，在光路中放置一片随机初始化的二值编码孔径，通过可微模型把掩膜参数直接嵌入网络前向传播，使得掩膜与重建网络在大规模合成-实拍混合数据集上通过训练而同时优化。由于神经网络与编码掩膜经过了大量数据的训练，这对实际的光场重建提供了隐式

的正则，即实际重建的光场也必须会满足一定的统计结构。图 10(b) 为实验结果，与最先进的方法在真实和合成光场数据集上相比，重建的光场不仅实现了更高的 PSNR(peak signal-to-noise ratio)，而且更好地保留了光场的视差结构，这对三维感知至关重要。图 10(c) 给出了不同采样数目下不同方法的性能，在各个采样数目下，Guo 等提出的方案始终保持更高的质量。

然而，端到端协同优化仍需预先设计好光学系统，而无法通过训练学习设计新型的光场编码采集方案。除了上述的经典孔径编码类压缩光场成像之外，近年来也有不少新型的光场采样方法。受无镜头 3D 成像领域的经典工作 DiffuserCam 的影响^[87]，Cai 等^[88]提出了无镜头光场成像，仅需一个廉价的散射片，就可以完成对 4D 光场的调制。如图 11(a) 所示，将入射光束分解为多个子光束，每个子光束照射散射片表面的小区域并形成子图像。图中， p_i 、 p_j 表示物空间不同的两个采样点。由于散射片表面的随机粗糙度分布，这些子图像彼此不同且不重叠，共同构成传感器上的完整图案。最关键的是，不同角度入射的光线被 diffuser 映射到传感器上的不同位置，从而实现了角度信息的空间编码。图 11(b) 为使用散射片对一个小型植物的光场成像。其中，左上角为原始数据，左下角为经过算法重建后得到的深度估计，右图则为重建后的光场对不同深度的后聚焦结果。

另一类思路是角度调制，利用多层或倾斜的

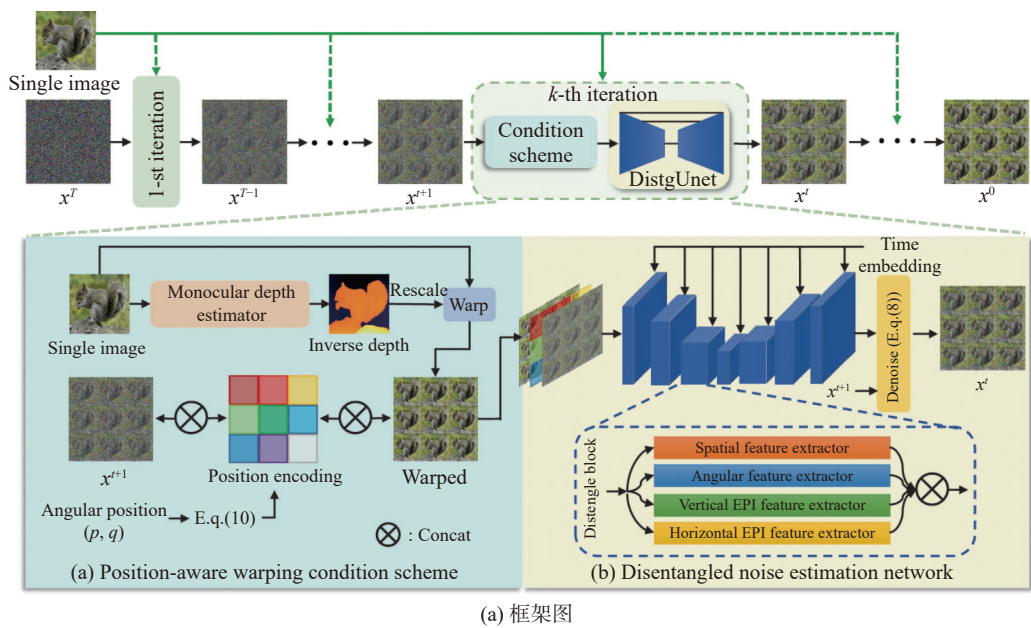


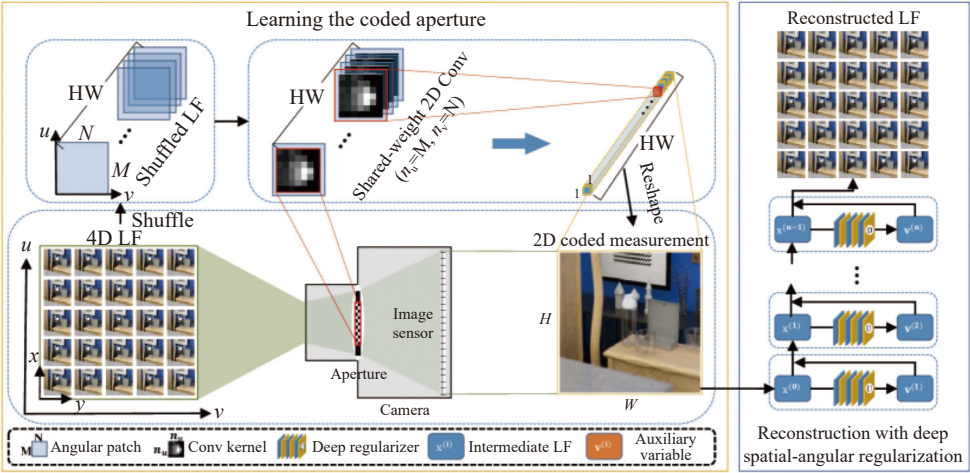
图 9 基于 diffusion 架构的压缩光场重构^[86]

Fig.9 Compressive light field reconstruction based on diffusion

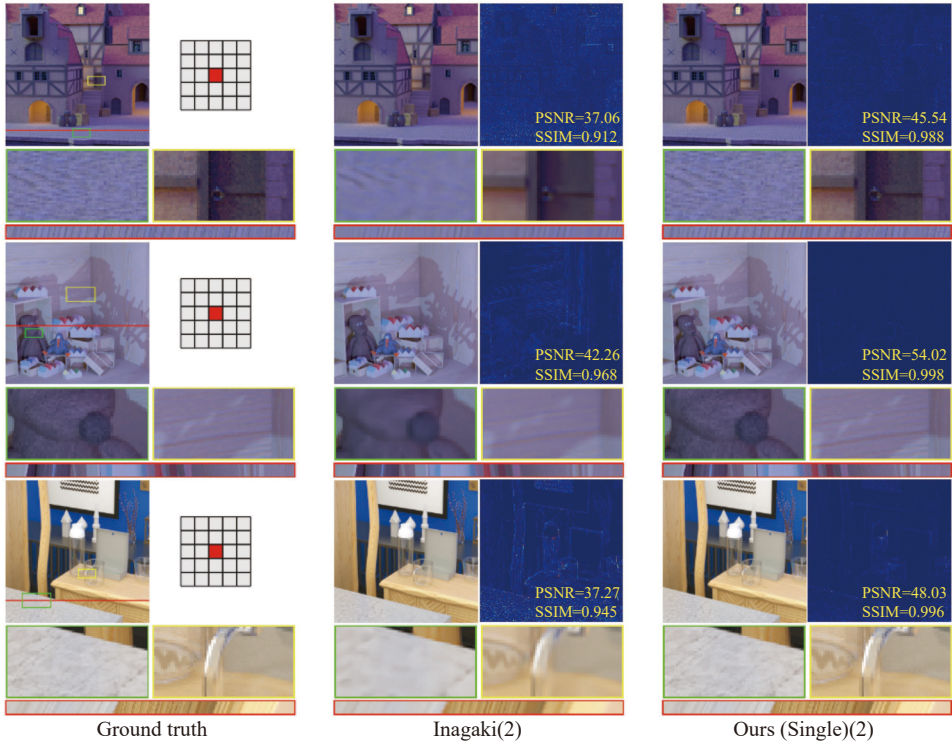
光学组件按角度对光线施加不同变换。例如柱面透镜阵列可以将物体在不同角度方向上的投影压缩到一维传感器上。2021 年，Feng 等^[89]提出的超高速光场断层成像 (LIFT) 系统中，采用了多个按不同倾斜角排列的微柱面透镜，将场景从不同视角的线积分投影同时引入条纹相机，实现快照式获取大尺度时变光场。如图 12 所示，通过 7 个角度的柱面透镜，LIFT 相机在单次曝光中采集到时间分辨的光场数据，从而以 30 Hz 的帧速实现非视线场景的三维重建。类似地，Feng 等^[90]在 2022 年提出的紧凑光场摄影 (CLIP) 框架融合了散射编码透镜和单像素/线阵探测等多种实现。例如，利用随机透镜产生随景深变化的广域点扩散函数，或在单像素相机中引入散焦透镜配合随机光学图案

进行测量。CLIP 通过非局部的光学混合采样，使每个子视图都包含整个场景的信息，大幅增强了对传感器缺陷和视线遮挡的鲁棒性。值得注意的是，编码光学方法往往需要针对特定光学设计开发相应重建算法。例如：对于单像素方案，由于每次仅获得图像的全局测量，需多次投影采样才能累积足够信息；而线阵传感器方案 (如 LIFT) 则结合断层重建算法和深度学习辅助，以在有限投影数下获得高质量结果。

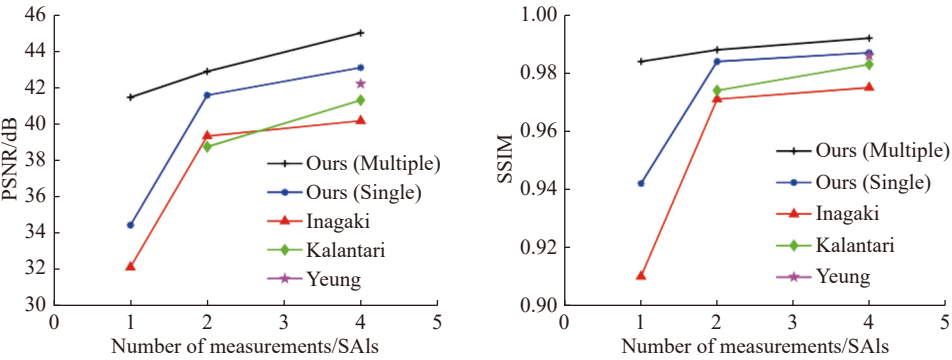
在压缩光场成像中，可能会在一些场景中不可避免地需要多次曝光采集光场。而多次曝光直接影响了成像的速度。为了解决这个问题，Habuchi 等^[91]提出了将时序编码与事件相机结合，在单次曝光内切换多个编码并用事件流捕获视差信息。



(a) 网络架构



(b) 效果图



(c) 与其他方案的PSNR/SSIM对比

图 10 采用神经网络的隐式正则代替稀疏先验时的可学习掩膜与深度重建网络的端到端优化^[46]

Fig.10 End-to-end optimization of learnable mask and depth reconstruction network with implicit regularization of neural network replacing sparse prior

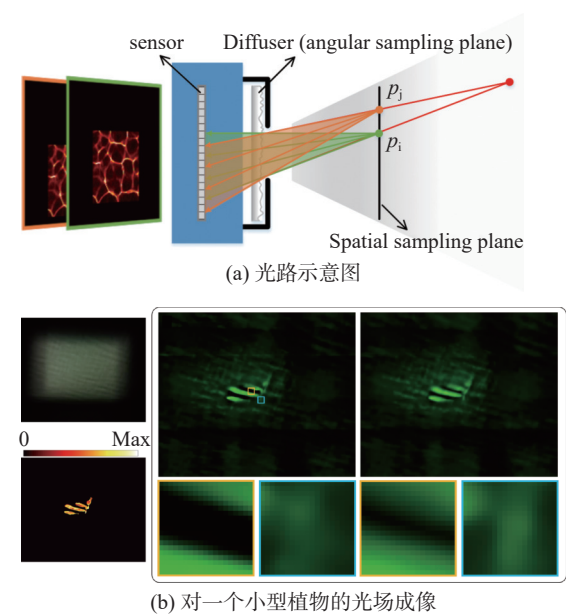


图 11 基于散射片的无镜头光场成像^[88]

Fig.11 Lensless light field imaging using a diffuser sheet

如图 13 所示，他们的方法不直接观察单个编码图像，而是将多个编码图像的总和作为单个图像帧，之后再用卷积神经网络重建出完整光场。图中： $L_{x,y,u,v}$ 表示 4D 光场； a_{uv} 表示不同孔径编码图案； $E_{x,y}$ 表示不同的事件； $I_{x,y}$ 表示不同的图像。由于目标场景是静态的，因此，事件完全由编码模式的变化引起。结果表明，这种方法可以通过

单次曝光比其他几种成像方法更准确地重建光场，以令人信服的视觉质量捕捉真实的 3D 场景。

Yun 等^[92]最近的工作还探索了捕获焦平面堆栈图进行光场重建。他们提出的 DDLFRM (denoising diffusion LF reconstruction model) 模型通过捕获少量不同对焦深度下的图片，依靠扩散模型的生成式先验，替代传统的稀疏、低秩正则，在欠采样条件下恢复高质量光场。如图 14(a) 所示：首先，建立了光场到焦平面图像的线性成像模型，并将其分解为可高效计算的系统矩阵；随后在扩散过程的每一步中，交替执行“生成式去噪(预测噪声/得分)”与“数据一致性校正(投影回投满足物理约束)”，逐步逼近真实光场分布。图 14(b) 展示了最终的成像结果以及与其他方法重构的光场效果对比，取得了可观的 PSNR 质量提升 (3 dB~7 dB)。

总体而言，人们提出了各种新型的光场编码采集方案以及对应的重构思路，使得光场成像系统具有更好的收集光场能力，可以在有限的采样中捕获更丰富的空间-角度信息。该方向的发展体现了光学硬件与计算模型深度融合的趋势。

表 1 列举对比了上文所提到的代表性工作的优缺点，结果表明，压缩光场成像的先验从早期人工设计的信号统计结构先验，如稀疏性，正逐

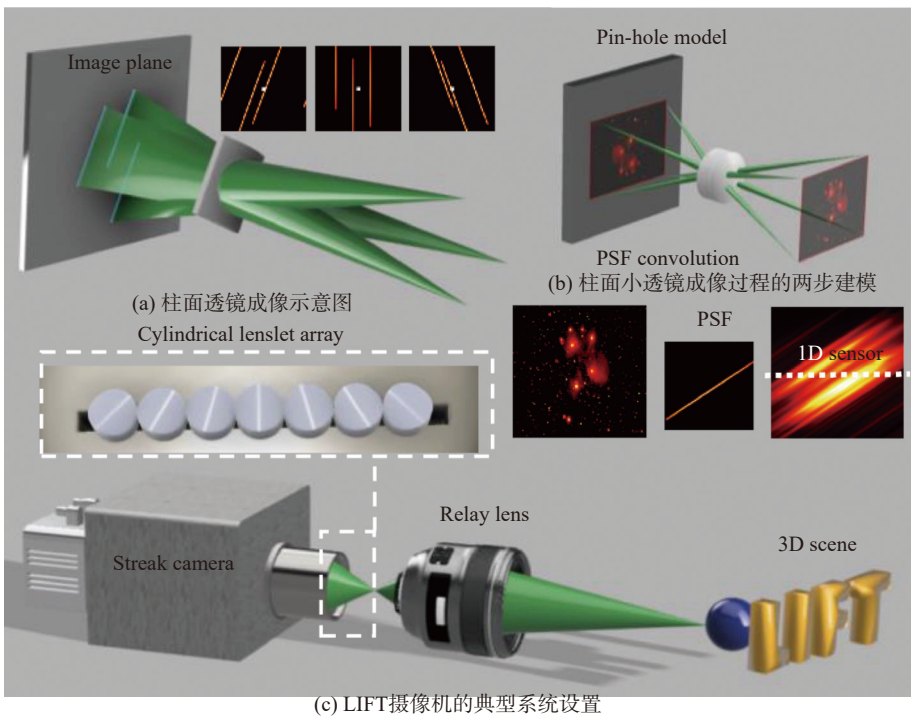


图 12 LIFT 光场编码示意图^[89]

Fig.12 Schematic diagram of LIFT light field encoding

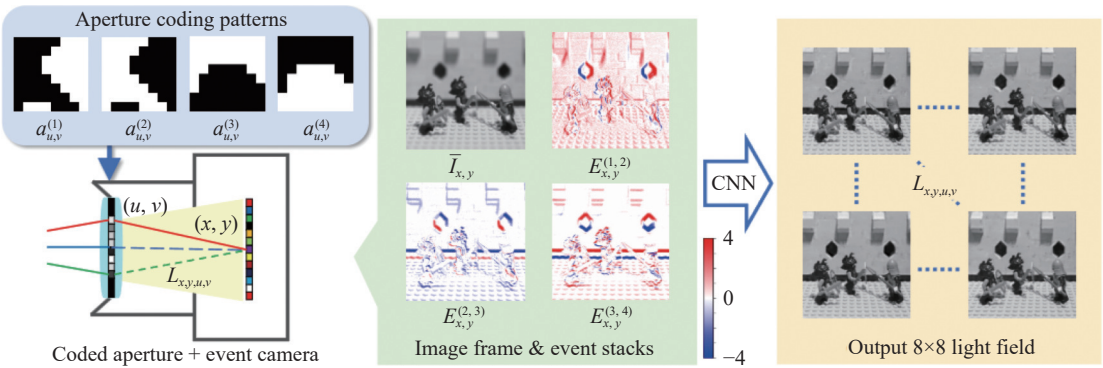


图 13 使用事件相机进行时序-孔径联合编码的光场成像^[91]

Fig.13 Light field imaging with temporal-aperture joint coding using an event camera

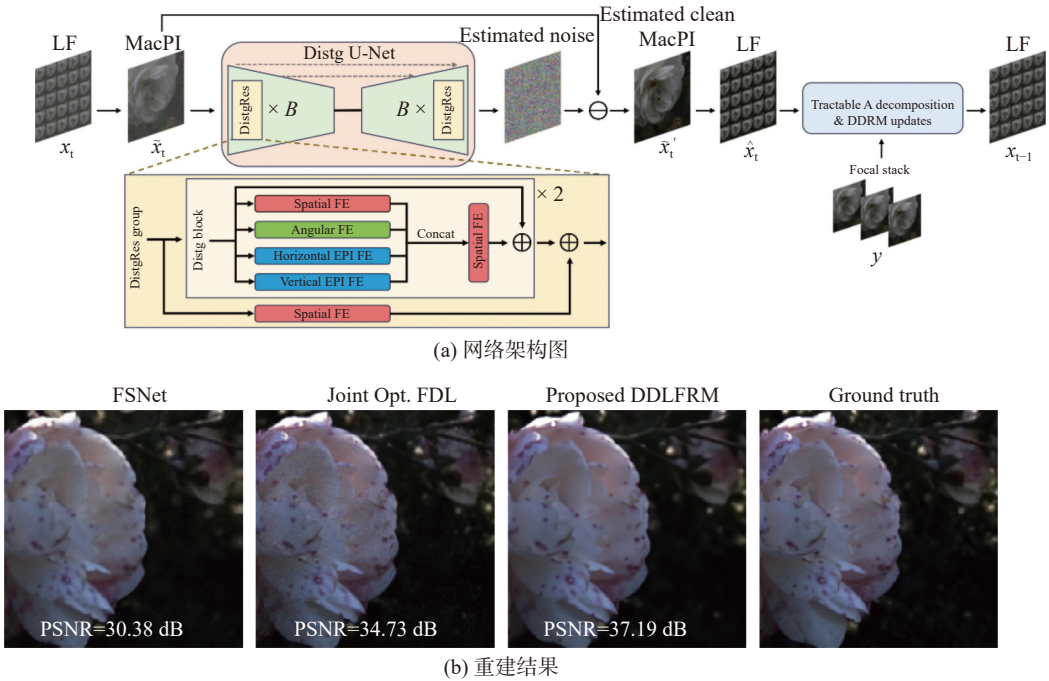


图 14 基于扩散模型的多深度聚焦堆叠图像光场重建^[92]

Fig.14 Light field reconstruction from focal-stack images captured at multiple depths using a diffusion model

表 1 不同压缩光场成像技术方案对比

Tab.1 Comparison of different compressive light field imaging techniques

技术路线	曝光次数	重建速度	复杂度	噪声鲁棒性	端到端	泛化性
稀疏重建/字典学习	单帧	慢	低	低	否	差
深度学习: CNN/Transformer	单帧	快	低	低	是	一般
深度学习: 扩散模型	单帧	慢	低	高	是	一般
调制掩膜+网络联合优化	单帧	快	低	低	是	一般
无镜头/散射片	多帧	慢	低	低	否	强
角度调制	多帧	慢	高	高	否	强
时序编码+事件相机	单帧	慢	高	低	否	强
焦平面堆栈	多帧	慢	低	高	是	一般

步发展到数据驱动的隐式先验。从系统上来看，人们研究出了各种各样的前向模型，最终都是为了增强光学系统收集光场的能力。从重建侧看，深度学习(含 Transformer、扩散模型)以端到端推

理显著加速，但对大规模数据与域一致性的依赖仍是关键瓶颈。综上，未来的优选路线是：在单帧/少帧采样下，采用简洁而可制造的编码光学，与物理可解释的端到端网络协同设计，从而在曝光次数、重建速度、硬件复杂度、噪声鲁棒性与泛化性之间达到更合理的统筹平衡。

4 应用

压缩光场成像以“光学编码 + 计算重建”为主线，成功将四维光场采样量由传统微透镜相机的百倍级压缩至单帧或少帧观测，在突破空间-角度折衷的同时保持了接近传感器原生的成像清晰度，为光场技术在三维感知、AR/MR、自适应光学及工业检测等多元场景中的落地奠定了坚实基础。

4.1 3D 成像

光场相机利用不同子孔径图片之间的视差可以提取物体的深度信息，从而实现 3D 成像。在 SLAM 领域、机器人等领域，基于光场相机的 3D 成像兼具“单镜头、长基线”的优势，通过对 4D 光场做基线视差分析，系统能直接输出绝对深度图，为对象在狭窄或低纹理环境中提供可靠的三维感知与避障基础。2020 年，Kaveti 等^[93]提出了光场的合成孔径成像。合成孔径技术是通过多视角数据融合的办法形成等效大孔径，从而在单次拍摄中实现像素级透视与深度估计。其关键优势在于抑制遮挡误差，并在同一步骤中联合估计深度与重聚焦图像，从而显著提升动态场景下的 SLAM 姿态精度与重定位成功率。这一技术展示了光场成像在 3D 场景几何重建与动态目标识别中的潜力。

4.2 AR/MR

光场成像为增强/混合现实 (AR/MR) 装备带来了突破性的“真三维”视觉体验。与传统立体或多焦段显示依赖两幅视图或少量深度平面的做法不同，光场显示在近眼位置重现完整的 4D 光线束，用户双眼可在连续景深内自由调焦，汇聚-调焦冲突显著降低，同时仍保持高分辨率、宽视场角与实时渲染能力。这一特性不仅缓解了长时间佩戴头显引起的视觉疲劳，还让数字内容能够在真实世界中以可信尺寸与正确遮挡关系“悬浮”，从而支持精准的空间定位、手势交互和近距操作等关键 AR/MR 场景。例如 Lanman 等^[94]提

出的 Near-Eye Light Field Displays 开创性地将微透镜阵列整合进轻量级头戴式显示器，通过 GPU 实时光场渲染在 10 mm 级厚度内同时提供正确的视差、调焦和视网膜散焦线索，系统原型验证了光场技术在消除汇聚-调焦冲突和减小头显体积方面的巨大潜力，奠定了后续光场 AR/MR 研究与产业化的基础。

4.3 自适应光学

在激光通信、天文观测等应用中，由于大气湍流等原因，光在传播过程中由于介质的不均匀而出现波前畸变。自适应光学 (AO) 的方法通过实时测量波前畸变进而高速补偿畸变，是实现高质量高分辨成像探测的关键技术之一。由于光场成像存在“空间-角度”的并采优势：同一张图像即可同时获得波前的相位梯度与振幅信息，无需多帧时序扫描或参考束干涉，使得其能够在一次曝光里解析复杂畸变，包括强湍流下的斑点、相位奇点以及大幅度波前跃迁，并实时反馈给可变形镜进行闭环校正。光场波前传感器不仅具备大动态范围，还对分束干涉引发的条纹、零强度点不敏感，可显著提升天文望远镜、自由空间光通信乃至高能激光传输系统中的畸变补偿效率和鲁棒性。如 Ko 等^[95]在 2017 年的工作，首次在实测大气湍流环境下系统比较了光场波前传感器与传统 S-H 方案的性能，证明了前者在深湍流 (Rytov 大于 1) 条件下可避免分裂斑点导致的重构失败，并在相同像元数量下实现更高的波前重建精度与动态范围，为今后“光场-AO”闭环控制的工程落地奠定了理论与实验基础。

4.4 器件缺陷检测与工业视觉

光场成像正迅速成为高速、非接触式器件缺陷检测的关键手段。在半导体晶圆、微型金属零件及 MEMS 批量生产线上，传统明 / 暗场或共焦检测方案往往只能给出二维纹理信息；而光场相机一次曝光即可同时获取被检组件的高分辨率 2D 纹理与亚毫级 3D 形貌，使系统能够在运动中的传送带上完成“纹理-几何”联合判别。得益于 4D 视差编码，算法可对高度反射或透明表面进行像素级深度分割，精准定位划痕、凹坑、翘曲等缺陷，并通过多视角冗余显著降低伪影率；其紧凑光机结构又易于嵌入生产线，实现 100 % 在线全检。2014 年，Weimer 等^[96]将光场相机集成至传送带上，利用单帧获取的 4D 数据在不足 50 ms 的

处理时间内同时输出缺陷纹理图与3D轮廓,成功实现了对微杯零件的全尺寸、全批次缺陷筛选,首次在微冲压金属零件产线上验证了光场相机的工业检测潜力。他们对比了光场方案与显微离线检测在检测速度、深度误差($<5\text{ }\mu\text{m}$)及伪检率上的差异,证明光场成像可在保持微米级精度的同时,将传统抽检流程提升为实时闭环质量控制。

5 未来挑战与展望

5.1 可微超构光学编码与算法协同设计

压缩光场成像的性能上限受限于光学编码与数值重建性能。传统手动设计掩膜往往难以兼顾光通量、调制稀疏性与制造可行性。超表面的快速发展以及液晶可编程器件等的快速发展,有望形成动态可调的高效4D光场编码,是下一代光学调控发展的方向。例如,通过电光调制信号调节液晶单元的相位延迟,以对成像物体进行频域动态调控。同时,完整光链的协同设计仍有巨大提升空间,联合优化光场编码方式与重建算法有望进一步提升压缩光场成像性能。

5.2 实时高分辨率重建

实时重建是压缩光场走向工业/消费级应用的关键。尽管已有GPU友好的渲染-重建管线,例如Chlubna等^[97]提出的轻量级全聚焦光场渲染在4K分辨率下已能25 fps运行,但端到端重建仍普遍受限于显存与带宽。Liu等^[98]提出的使用光场探针进行实时场景重建,提出在多GPU交换式管线中利用视差一致约束,实现了更高吞吐量。但是,光场的4D数据张量巨大,需要稀疏表示、混合精度与块迭代策略。并且,动态场景下的多帧稳像与实时深度估计尚缺一解决方案。

5.3 基于神经网络重建光场的鲁棒泛化与跨域适应

当前压缩光场重建网络多数在合成数据集上训练,实拍迁移性能有限。“域偏移(domain shift)”带来噪声分布、光谱响应差异,严重影响视差一致性。在镜头污渍、编码片翘曲、CMOS噪声等情况下保持重建质量也是一个非常重要的问题。Hu等^[99]提出的基于空间-角度密集网络的高效光场重建通过密集跳连缓解了域不一致,但仍依赖大规模标注。Wang等^[100]提出的MDANet多阶段域适应框架则在低光照场景中验证了无标

注跨域重建的可行性,为今后神经网络重建光场提供思路。

5.4 多模态融合与应用扩展

压缩光场成像正与飞行时间成像(ToF)、非视距成像(NLOS)、全息显微等方向加速融合。Feng等^[90]展示了利用稀疏传感阵列与解析解重建,实现紧凑3D视觉模组;Peng等^[101]则把高速编码与显微荧光结合,在活体观察中实现毫秒级3D成像。未来的光场重建可能会与各种多模态信息共同压缩,提升信息密度。

5.5 总结与展望

光场成像通过同时捕捉场景中的完整4D光场,同时记录下光场的角度与空间信息,为多尺度、多场景三维感知打开了新的想象空间,并在各种应用场合中发挥出不可替代的作用。而压缩光场重建凭借“压缩采样获取高维场景信息、后端算法解码重构”的特点,为光场成像的分辨率与质量等各种维度上的提升提供了可能。未来,我们相信压缩光场成像技术会进一步商业化与实用化,在各种领域中更加深刻地影响人们的工作与生活。

参考文献:

- [1] LEVOY M, NG R, ADAMS A, et al. Light field microscopy[J]. *ACM Transactions on Graphics*, 2006, 25(3): 924–934.
- [2] NG R. Fourier slice photography[C]//SIGGRAPH'05: ACM SIGGRAPH 2005 Papers. Los Angeles: ACM, 2005: 735–744.
- [3] LEVOY M, HANRAHAN P. Light field rendering[C]//Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques. New Orleans: ACM, 1996: 31–42.
- [4] VAISH V, GARG G, TALVALA E, et al. Synthetic aperture focusing using a shear-warp factorization of the viewing transform[C]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). San Diego: IEEE, 2005: 129.
- [5] ISAKSEN A, MCMILLAN L, GORTLER S J. Dynamically reparameterized light fields[C]//Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. New Orleans: ACM, 2000: 297–306.
- [6] LEVOY M, CHEN B, VAISH V, et al. Synthetic aperture confocal imaging[J]. *ACM Transactions on Graphics*, 2004, 23(3): 825–834.

- [7] GORTLER S J, GRZESZCZUK R, SZELISKI R, et al. The lumigraph[C]//Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques. New Orleans: ACM, 1996: 43–54.
- [8] WETZSTEIN G, LANMAN D, HIRSCH M, et al. Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting[J]. *ACM Transactions on Graphics*, 2012, 31(4): 80.
- [9] 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: IEEE, 2018: 4530–4538.
- [10] 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 (CVPR). Boston: IEEE, 2015: 1547–1555.
- [11] 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.
- [12] WANNER S, GOLDLUECKE B. Globally consistent depth labeling of 4D light fields[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2012: 41–48.
- [13] HEDMAN P, PHILIP J, PRICE T, et al. Deep blending for free-viewpoint image-based rendering[J]. *ACM Transactions on Graphics*, 2018, 37(6): 257.
- [14] KALANTARI N K, WANG T C, RAMAMOORTHY R. Learning-based view synthesis for light field cameras[J]. *ACM Transactions on Graphics*, 2016, 35(6): 193.
- [15] LIPPMANN G. La photographie integrale[J]. *Comptes Rendus, Academie des Sciences*, 1908, 146: 446–451.
- [16] GERSHUN A. The light field[J]. *Journal of Mathematics and Physics*, 1939, 18(1/4): 51–151.
- [17] ADELSON E H, BERGEN J R. The plenoptic function and the elements of early vision[M]//LANDY M, MOVSHON J A. *Computational Models of Visual Processing*. Cambridge: The MIT Press, 1991: 3–20.
- [18] MAHMOUDPOUR S, PAGLIARI C, SCHELKENS P. Learning-based light field imaging: an overview[J]. *EURASIP Journal on Image and Video Processing*, 2024, 2024(1): 12.
- [19] 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.
- [20] BROXTON M, GROSENICK L, YANG S, et al. Wave optics theory and 3-D deconvolution for the light field microscope[J]. *Optics Express*, 2013, 21(21): 25418–25439.
- [21] 张润南, 周宁, 周子豪, 等. 光场表征及其分辨率提升技术: 文献综述及最新进展 [J]. *红外与激光工程*, 2024, 53(9): 20240347.
- [22] LUMSDAINE A, GEORGIEV T. The focused plenoptic camera[C]//2009 IEEE International Conference on Computational Photography (ICCP). San Francisco: IEEE, 2009: 1–8.
- [23] 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.
- [24] NG R, LEVOY M, BRÉDIF M, et al. Light field photography with a hand-held plenoptic camera[R]. Stanford: Stanford University, 2005.
- [25] PERWAß C, WIETZKE L. Single lens 3D-camera with extended depth-of-field[C]//Proceedings of SPIE 8291, Human Vision and Electronic Imaging XVII. Burlingame: SPIE, 2012: 829108.
- [26] 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: IEEE, 2013: 1027–1034.
- [27] 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.
- [28] CONTI C, SOARES L D, NUNES P. Dense light field coding: a survey[J]. *IEEE Access*, 2020, 8: 49244–49284.
- [29] ZHOU S Y, ZHU T Q, SHI K L, et al. Review of light field technologies[J]. *Visual Computing for Industry, Biomedicine, and Art*, 2021, 4(1): 29.
- [30] LIANG C K, RAMAMOORTHY R. A light transport framework for lenslet light field cameras[J]. *ACM Transactions on Graphics*, 2015, 34(2): 16.
- [31] 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.
- [32] BABACAN S D, ANSORGE R, LUESSI M, et al. Compressive light field sensing[J]. *IEEE Transactions on Image Processing*, 2012, 21(12): 4746–4757.
- [33] DAVIS A, LEVOY M, DURAND F. Unstructured light fields[J]. *Computer Graphics Forum*, 2012, 31(2): 305–314.
- [34] 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. Pisa: ACM, 2002: 77–86.

- [35] VAISH V, WILBURN B, JOSHI N, et al. Using plane + parallax for calibrating dense camera arrays[C]//Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. Washington: IEEE, 2004: 2–9.
- [36] 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. Washington: IEEE, 2004: 294–301.
- [37] PEI Z, LI Y W, MA M, et al. Occluded-object 3D reconstruction using camera array synthetic aperture imaging[J]. *Sensors*, 2019, 19(3): 607.
- [38] VAISH V, LEVOY M, SZELISKI R, et al. Reconstructing occluded surfaces using synthetic apertures: stereo, focus and robust measures[C]//2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR' 06). New York: IEEE, 2006: 2331–2338.
- [39] WILBURN B, JOSHI N, VAISH V, et al. High performance imaging using large camera arrays[J]. *ACM Transactions on Graphics (TOG)*, 2005, 24(3): 765–776.
- [40] LIANG C K, SHIH Y C, CHEN H H. Light field analysis for modeling image formation[J]. *IEEE Transactions on Image Processing*, 2011, 20(2): 446–460.
- [41] GEORGIEV T G, LUMSDAINE A. Focused plenoptic camera and rendering[J]. *Journal of Electronic Imaging*, 2010, 19(2): 021106.
- [42] WU J M, GUO Y D, DENG C, et al. An integrated imaging sensor for aberration-corrected 3D photography[J]. *Nature*, 2022, 612(7938): 62–71.
- [43] ASHOK A, NEIFELD M A. Compressive light field imaging[C]//Proceedings of SPIE 7690, Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV. Orlando: SPIE, 2010: 76900Q.
- [44] 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–es.
- [45] 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.
- [46] GUO M T, HOU J H, JIN J, et al. Deep spatial-angular regularization for compressive light field reconstruction over coded apertures[C]//Proceedings of the 16th European Conference on Computer Vision. Glasgow: Springer, 2020: 278–294.
- [47] 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.
- [48] GUPTA M, JAUHARI A, KULKARNI K, et al. Compressive light field reconstructions using deep learning[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Honolulu: IEEE, 2017: 1277–1286.
- [49] MIANDJI E, HAJISHARIF S, UNGER J. A unified framework for compression and compressed sensing of light fields and light field videos[J]. *ACM Transactions on Graphics*, 2019, 38(3): 23.
- [50] DONOHO D L, MALEKI A, MONTANARI A. Message-passing algorithms for compressed sensing[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, 106(45): 18914–18919.
- [51] BECK A, TEOULLE M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems[J]. *SIAM Journal on Imaging Sciences*, 2009, 2(1): 183–202.
- [52] DAI W, MILENKOVIC O. Subspace pursuit for compressive sensing signal reconstruction[J]. *IEEE Transactions on Information Theory*, 2009, 55(5): 2230–2249.
- [53] NEEDELL D, TROPP J A. CoSaMP: iterative signal recovery from incomplete and inaccurate samples[J]. *Applied and Computational Harmonic Analysis*, 2009, 26(3): 301–321.
- [54] TROPP J A, GILBERT A C. Signal recovery from random measurements via orthogonal matching pursuit[J]. *IEEE Transactions on Information Theory*, 2007, 53(12): 4655–4666.
- [55] DONOHO D L. Compressed sensing[J]. *IEEE Transactions on Information Theory*, 2006, 52(4): 1289–1306.
- [56] CHEN S S, DONOHO D L, SAUNDERS M A. Atomic decomposition by basis pursuit[J]. *SIAM Review*, 2001, 43(1): 129–159.
- [57] CANDES E J, ROMBERG J, TAO T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information[J]. *IEEE Transactions on Information Theory*, 2006, 52(2): 489–509.
- [58] MARQUEZ M, RUEDA-CHACON H, ARGUELLO H. Compressive spectral light field image reconstruction via online tensor representation[J]. *IEEE Transactions on Image Processing*, 2020, 29: 3558–3568.
- [59] ASHOK A, NEIFELD M A. Compressive light field imaging with weighted random projections[C]//Proceedings of SPIE 8165, Unconventional Imaging,

- Wavefront Sensing, and Adaptive Coded Aperture Imaging and Non-Imaging Sensor Systems. San Diego: SPIE, 2011: 816519.
- [60] 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.
- [61] HOSSEINI KAMAL M, HESHMAT B, RASKAR R, et al. Tensor low-rank and sparse light field photography[J]. *Computer Vision and Image Understanding*, 2016, 145: 172–181.
- [62] BLOCKER C J, CHUN Y, FESSLER J A. Low-rank plus sparse tensor models for light-field reconstruction from focal stack data[C]//2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP). Arishti Village: IEEE, 2018: 1–5.
- [63] LIU C, QIU J, JIANG M. Light field reconstruction from projection modeling of focal stack[J]. *Optics Express*, 2017, 25(10): 11377–11388.
- [64] PÉREZ F, PÉREZ A, RODRÍGUEZ M, et al. Lightfield recovery from its focal stack[J]. *Journal of Mathematical Imaging and Vision*, 2016, 56(3): 573–590.
- [65] LE PENDU M, JIANG X R, GUILLEMOT C. Light field inpainting propagation via low rank matrix completion[J]. *IEEE Transactions on Image Processing*, 2018, 27(4): 1981–1993.
- [66] NICKOLLS J, BUCK I, GARLAND M, et al. Scalable parallel programming with CUDA: is CUDA the parallel programming model that application developers have been waiting for?[J]. *Queue*, 2008, 6(2): 40–53.
- [67] JOUPPI N P, YOUNG C, PATIL N, et al. In-datacenter performance analysis of a tensor processing unit[C]//Proceedings of the 44th Annual International Symposium on Computer Architecture. Toronto: ACM, 2017: 1–12.
- [68] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. *Nature*, 2015, 521(7553): 436–444.
- [69] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84–90.
- [70] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 770–778.
- [71] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: IEEE, 2017: 6000–6010.
- [72] CHEN T, KORNBLITH S, NOROUZI M, et al. A simple framework for contrastive learning of visual representations[C]//Proceedings of the 37th International Conference on Machine Learning. New York: PMLR, 2020: 149.
- [73] HE K M, FAN H Q, WU Y X, et al. Momentum contrast for unsupervised visual representation learning[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020: 9726–9735.
- [74] HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models[C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver: Curran Associates, Inc., 2020: 574.
- [75] KARRAS T, AITTALA M, LAINE S, et al. Elucidating the design space of diffusion-based generative models[C]//Proceedings of the 36th International Conference on Neural Information Processing Systems. New Orleans: Curran Associates Inc., 2022: 1926.
- [76] SONG Y, SOHL-DICKSTEIN J, KINGMA D P, et al. Score-based generative modeling through stochastic differential equations[EB/OL]. (2020-11-26). [2025-07-30]. <https://doi.org/10.48550/arXiv.2011.13456>.
- [77] VINCENT P. A connection between score matching and denoising autoencoders[J]. *Neural Computation*, 2011, 23(7): 1661–1674.
- [78] SONG Y. Learning to generate data by estimating gradients of the data distribution[D]. Stanford: Stanford University, 2022.
- [79] LIPMAN Y, CHEN R T Q, BEN-HAMU H, et al. Flow matching for generative modeling[C]//Proceedings of the 11th International Conference on Learning Representations. Kigali: OpenReview. net, 2023.
- [80] SUTSKEVER I, VINYALS O, LE Q V. Sequence to sequence learning with neural networks[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal: ACM, 2014: 3104–3112.
- [81] AMODEI D, ANANTHANARAYANAN S, ANUBHAI R, et al. Deep speech 2: end-to-end speech recognition in English and mandarin[C]//Proceedings of the 33rd International Conference on International Conference on Machine Learning. New York: PMLR, 2016: 173–182.
- [82] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision[C]//Proceedings of the 38th International Conference on Machine Learning. New York: PMLR, 2021: 8748–8763.

- [83] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 9992–10002.
- [84] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale[EB/OL]. (2020-10-22)[2025-07-30]. <https://doi.org/10.48550/arXiv.2010.11929>.
- [85] HUA X Y, WANG M H, SU B N, et al. Densely sampled light field reconstruction with transformers[J]. *Journal of Electronic Imaging*, 2023, 32(3): 033025.
- [86] GAO R S, LIU Y T, XIAO Z Y, et al. Diffusion-based light field synthesis[C]//DEL BUE A, CANTON C, PONT-TUSET J, et al. International Conference Computer Vision. Milan: Springer, 2024: 1–19.
- [87] ANTIPA N, KUO G, HECKEL R, et al. DiffuserCam: lensless single-exposure 3D imaging[J]. *Optica*, 2018, 5(1): 1–9.
- [88] CAI Z W, CHEN J W, PEDRINI G, et al. Lensless light-field imaging through diffuser encoding[J]. *Light: Science & Applications*, 2020, 9(1): 143.
- [89] FENG X H, GAO L. Ultrafast light field tomography for snapshot transient and non-line-of-sight imaging[J]. *Nature Communications*, 2021, 12(1): 2179.
- [90] FENG X H, MA Y Y, GAO L. Compact light field photography towards versatile three-dimensional vision[J]. *Nature Communications*, 2022, 13(1): 3333.
- [91] HABUCHI S, TAKAHASHI K, TSUTAKE C, et al. Time-efficient light-field acquisition using coded aperture and events[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2024: 24923–24933.
- [92] YUN H S, CHUN I Y. Improving light field reconstruction from limited focal stack using diffusion models[C]//2024 IEEE 34th International Workshop on Machine Learning for Signal Processing (MLSP). London: IEEE, 2024: 1–6.
- [93] KAVETI P, SINGH H. A light field front-end for robust SLAM in dynamic environments[EB/OL]. (2020-12-19). <https://doi.org/10.48550/arXiv.2012.10714>.
- [94] LANMAN D, LUEBKE D. Near-eye light field displays[J]. *ACM Transactions on Graphics*, 2013, 32(6): 220.
- [95] KO J, DAVIS C C. Comparison of the plenoptic sensor and the Shack–Hartmann sensor[J]. *Applied Optics*, 2017, 56(13): 3689–3698.
- [96] WEIMER D, THAMER H, FELLMANN C, et al. Towards 100% in-situ 2D/3D quality inspection of metallic micro components using plenoptic cameras[J]. *Procedia CIRP*, 2014, 17: 847–852.
- [97] CHLUBNA T, MILET T, ZEMČÍK P. Lightweight all-focused light field rendering[J]. *Computer Vision and Image Understanding*, 2024, 244: 104031.
- [98] LIU Y R, NOWROUZEZAHRI D, MCGUIRE M. Real-time scene reconstruction using light field probes[EB/OL]. (2025-07-19). <https://doi.org/10.48550/arXiv.2507.14624>.
- [99] HU Z X, YEUNG H W F, CHEN X M, et al. Efficient light field reconstruction via spatio-angular dense network[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 5014314.
- [100] WANG J H, HE Y F, LI K, et al. MDANet: a multi-stage domain adaptation framework for generalizable low-light image enhancement[J]. *Neurocomputing*, 2025, 627: 129572.
- [101] PENG J Z, HUANG S Y, LI J P, et al. Differential high-speed aperture-coding light field microscopy for dynamic sample observation with enhanced contrast[J]. *Advanced Imaging*, 2024, 1(3): 031002.

(编辑: 丁红艺)