

四维光场表达模型综述

李亚宁, 王雪, 周果清, 王庆*

西北工业大学计算机学院, 陕西 西安 710072

摘要 光场成像与处理是计算摄像学领域最具潜力的发展方向之一。在获得 4D 光场的基础上, 构建简洁且高效的光场表达模型是光场理论的重要研究内容, 也是光场技术迈向应用的关键。本文从 4D 光场的深度、纹理、几何、尺度和频谱等角度出发对光场表达模型进行解析, 分别介绍子孔径图像、对极平面图、聚焦栈、超像素、多平面图像、超图、傅里叶切片、傅里叶视差层和对极聚焦谱等相关光场表达的特性, 分析光场表达模型在光场重建、深度估计和光场编辑等不同应用场景的适用条件, 归纳总结各种光场表达模型的优缺点及其潜在的应用方向。

关键词 成像系统; 光场表达; 计算摄像学; 傅里叶切片; 光场重建

中图分类号 TP37

文献标志码 A

doi: 10.3788/LOP202158.1811012

Overview of 4D Light Field Representation

Li Yaning, Wang Xue, Zhou Guoqing, Wang Qing*

School of Computer Science, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China

Abstract Light field imaging and processing is one of the most potential directions in the field of computational photography. Constructing concise and efficient representations of light field is the theoretical basis and key to promote the applications of light field. Based on the depth, texture, geometry, scale, and spectrum of light field, this paper analyzes characteristics of typical light field representations, including sub-aperture image, epipolar plane image, focal stack, super-pixel, multiplane images, hypergraph, Fourier slice, Fourier disparity layer, and epipolar focus spectrum. This paper also summarizes the impacts of different representations for light field applications, such as reconstruction, depth estimation, editing and so on. Finally, the advantages, disadvantages, and potential applications of various representations are summarized.

Key words imaging systems; light field representation; computational photography; Fourier slice; light field reconstruction

OCIS codes 110.1758; 110.2960; 100.2000

1 引言

光线是人类感知世界的主要媒介之一, 而如何记录和描述光线是机器视觉的基本理论问题。传统的成像技术只能记录光线的位置信息, 而光场成像技术可以同时记录光线的位置及角度信息^[1-2], 并且能够实现多种新颖的成像效果, 如数字重聚焦、多视角成像^[3]和景深扩展等^[4], 使其在高动态成像^[5]、层

析成像^[6-7]、增强现实^[8]、黑洞成像^[9]和细胞神经活动观测^[10-11]等方面具有广泛的应用前景, 因此光场成像是计算摄像学^[11]最具潜力的发展方向之一。

4D 光场数据中特有的角度信息为场景理解提供了强大支撑, 拓展了机器视觉的适用范围, 提高了视觉问题的求解精度。在光场深度估计^[12-13]应用方面, 区别于立体视觉匹配方法, 基于光场的方法能够融合包括聚焦/离焦、聚焦对称性、多视立体匹配和

收稿日期: 2021-05-31; 修回日期: 2021-06-09; 录用日期: 2021-06-15

基金项目: 国家自然科学基金(62031023, 61801396)

通信作者: *qwang@nwpu.edu.cn

连续域光强分布属性^[14]等多种深度线索,能够更好地估计遮挡边缘的深度,从而提高深度估计算法的精确度与鲁棒性。受到带宽积的限制,现有的光场成像设备难以同时获取高空间和高角度分辨率的光场,因此有必要研究高分辨率光场的重建方法,而现有方法^[15]主要包括依赖深度、不依赖深度及半依赖深度方法。第一类依赖深度方法首先估计光场深度,然后根据深度信息将现有的视图变换到新视图。第二类不依赖深度方法是直接对不同视点的纹理分布或其频域特性进行建模,可以获得高质量的视点超分辨率结果。第三类方法则受益于深度学习技术,其核心是构造多平面图像(Multiplane Images, MPI),即场景中不同深度层物体的纹理表达,在生成新视点图像的过程中,只需要将不同层的纹理进行投影即可。与传统二维图像相比,密集角度采样赋予光场亚像素精度的超分辨率重建能力,能够有效提升重建结果的清晰度。在显微成像方面,景深是限制数字显微在大纵深视场中应用的主要问题,4D 光场的多视点采样能够有效扩展景深,为显微三维观测^[11]以及裸眼三维显示^[16-17]等应用带来新的机遇。

然而,光场的巨大数据冗余为采集^[18]、存储和传输^[19]带来了极大挑战,因此研究人员在 4D 光场基本表达模型的基础上,通过分析其维度、结构、纹理、语义、深度以及频域等特性,构建了多种简洁高效的光场表达模型。不同的光场表达模型在光场重建和深度估计等应用方面分别展示了各自优越的性能。例如,基于光场角度维特性的表达能够更好地保持视点一致性;基于纹理和深度特性的表达包含光场的主要信息分布,在去除冗余信息的同时能够更好地处理光场重建和深度估计所面临的遮挡问题;基于结构和语义信息的表达能够构建场景内部的结构关联;基于频域的表达不需要显式的深度信

息,能够避免光场重建过程中深度歧义性所导致的偏差。

综上所述,简洁高效的光场表达模型是光场成像与处理技术的研究重点之一。随着光场成像技术的快速发展,研究人员提出了多种光场表达模型,为了方便本领域研究人员全面了解光场表达模型,本文将现有的 4D 光场表达模型及其相关应用进行总结归纳。

2 光场表达模型

为了完备描述空间光线的基本属性,麻省理工学院的 Adelson 教授^[20]于 1991 年提出了 7D 全光函数 $L(x, y, z, \theta, \phi, \lambda, t)$, 该函数描述了空间中的任意位置 (x, y, z) 、任意角度 (θ, ϕ) 、任意波长 λ 和任意时刻 t 的光线属性。针对不同的应用需求,可以对全光函数进行降维采样和表达。在假设光线强度不随时间衰减的前提下,可约去波长维 λ 及时间维 t , 从而得到包含空间维和角度维采样的 4D 光场。在此基础上,研究人员提出了基于球面点对、点球以及点切的光场表达模型^[21]、基于双柱面的全景光场表达模型^[22], 以及双平面光场表达模型^[23-24]。其中双平面光场表达模型由于具有模型简单和数据存储方便等优势,使其在学术界和工业界得到了广泛的应用。因此,本文主要从空域和频域两个方面对不同数据特性的双平面光场表达模型进行介绍。

2.1 空域光场表达模型

2.1.1 纹理-场景表达模型——4D 光场基本表达 (子孔径图像)

1996 年,Levoy 等^[23]和 Gortler 等^[25]提出了 4D 光场的基本表达模型 $L(u, v, x, y)$ 。根据该表达模型,双平面可同时记录空间中光线的角度和位置信息。如图 1(a)所示,视点平面 (u, v) 位于 $Z=0$ 的平面上,记录光线的方向信息。图像平

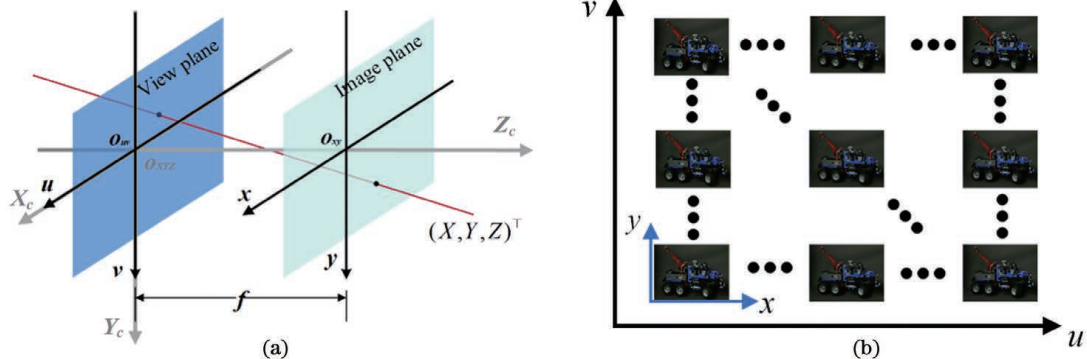


图 1 4D 光场的基本表达^[23]。(a) 双平面参数化表达;(b) 子孔径图像阵列

Fig. 1 Basic representation of 4D light field^[23]. (a) Parametric representation of two-parallel plane; (b) sub-aperture image array

面 (x, y) 与视点平面平行, 位于相机坐标系 $Z = f$ 的平面上, 记录光线的位置信息, 其中 f 为双平面间的距离。对于空间中点 (X, Y, Z) 发出的任一光线, 仅需知道其与视点平面和图像平面的交点 (u, v) 和 (x, y) , 便可唯一确定。

当固定角度坐标时, 每个视点可看作传统小孔成像模型, 但主点位置有偏移。空间中三维点 $[X \ Y \ Z]^T$ 发出的一条光线被双平面所记录, 设其交点分别为 (u, v) 与 (x, y) , 则 4D 光场在空间维和角度维的采样模型可表示为^[26-27]

$$\lambda_{xy} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & u & -fu \\ 0 & f & v & -fv \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (1)$$

$$\lambda_{uv} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & -x & 0 \\ 0 & f & -y & 0 \\ 0 & 0 & 1 & -f \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (2)$$

式中: λ_{xy} 和 λ_{uv} 为深度相关的缩放因子, $\lambda_{xy} = Z$, $\lambda_{uv} = Z - f$ 。考虑到场景的整体结构和纹理信息, 可以使用 2D 图像阵列来表示 4D 光场, 也就是通常所说的子孔径图像 $\mathbf{I}_{(u,v)} \in \mathbf{R}^{H \times W}$, 如图 1(b) 所示, 其中 H 和 W 分别为子孔径图像的高和宽, 而且每个子孔径图像对应一组固定的 (u, v) 坐标。由图 1 可见, 子孔径图像包含了一个视点的场景纹理信息, 因此其数量即为光场角度的采样数。但是子孔径图像序列本质上是传统的 2D 图像阵列, 无法直接表达不同子孔径图像之间的关联关系。

由于不同子孔径图像之间存在视差关系, 因此常被用于场景深度估计以及视点重建。在深度估计的过程中, 不同于立体匹配中只能获取两个视点信息, 子孔径图像能够提供更多的约束, 可以得到更精确的深度估计结果。Yu 等^[28]探索了空间点在 4D 光场中的超平面结构, 通过在图像分割框架中引入平面约束可以提升立体匹配的精度。Jeon 等^[29]定义子孔径图像间的相似性度量, 获得了更精确的深度估计结果。Wang 等^[30]提出了一个遮挡模型, 基于该模型能够更准确地对图像边缘进行深度估计。在此基础上, Zhu 等^[31]提出了一个完整的光场遮挡一致性模型, 用于处理多遮挡情况下的深度估计。

光场视点的重建是基于图像渲染 (Image-based Rendering, IBR)^[32]问题的特例, 即输入视点与拟重建视点均位于规则 2D 网格上的 IBR。因此, 当输入为子孔径图像时, 传统的基于深度估计的 IBR 方

法^[33-37]可以被直接应用于视点重建^[38-39]。Wanner 等^[14]提出了一种基于深度估计的光场视点合成方法。Bishop 等^[40]利用了光场多视角图像之间的冗余信息对光场图像进行超分辨率重建。传统的基于深度的视点合成方法过度依赖于深度估计, 导致其在无纹理和遮挡区域的重建效果不佳。因此, 卷积神经网络 (Convolutional Neural Network, CNN)^[41]逐渐被研究人员应用于光场视点的重建。以稀疏视点的 4D 光场作为网络输入, Wang 等^[42]提出了一种直接对稀疏光场进行角度域重建的伪 4D 神经网络。Yeung 等^[43]对稀疏 4D 光场所采样的粗特性进行分析, 提出了一种角度维和空间维交替卷积的光场重建方法, 提高了视点重建效率。Flynn 等^[44]基于学习方法提出了从一组宽基线视图序列合成新视图的方法。Kalantari 等^[45]使用两个 CNN 同时估计深度和纹理, 并通过最小化合成视图与真实图像之间的误差实现了新视点的合成。Yoon 等^[46]通过数据驱动的学习方法实现了相邻视图的超分辨率重建。赵圆圆等^[47]首先通过多尺度特征提取模块探索光场中子孔径图像的固有结构信息, 然后使用融合模块对提取到的纹理信息进行融合增强, 最后使用上采样模块实现光场图像阵列的超分辨率成像。在显微成像方面, 杨叶凡等^[48]对光场视图之间的互补信息进行正则化平滑处理, 得到了高分辨率的光场显微成像结果。光场成像还可用于三维测量, 伍俊龙等^[49]利用了光场子孔径图像之间的细微纹理差异, 准确获取弱纹理区域和精细结构区域的视差信息以复原场景的三维结构, 从而得到准确的三维测量结果。

2.1.2 纹理-几何表达模型——对极平面图 (EPI)

4D 光场的另一个基本特点是角度维采样的结构特性^[23], 结构化的视点采样使得不同视点之间的像素排布具有特定的规律性。从对极平面图 (Epipolar Plane Image, EPI) 中可以明显观察到光线在不同视点的分布关联^[3]。对于双平面光场, 通过固定 4D 光场的一个空间维和对应的角度维可得到其 EPI, 如图 2(a) 所示。

双平面光场中角度维和空间维之间的关系以及 EPI 的数学表达式分别为^[50]

$$\begin{cases} u = X - \frac{B}{f} \frac{x}{d^{(x,y)}} \\ v = Y - \frac{B}{f} \frac{y}{d^{(x,y)}} \end{cases}, \quad (3)$$

$$F_{\text{EPI}}(u, x) = F_{\text{EPI}}\left(X - \frac{B}{f} \frac{x}{d^{(x,y)}}, x\right), \quad (4)$$

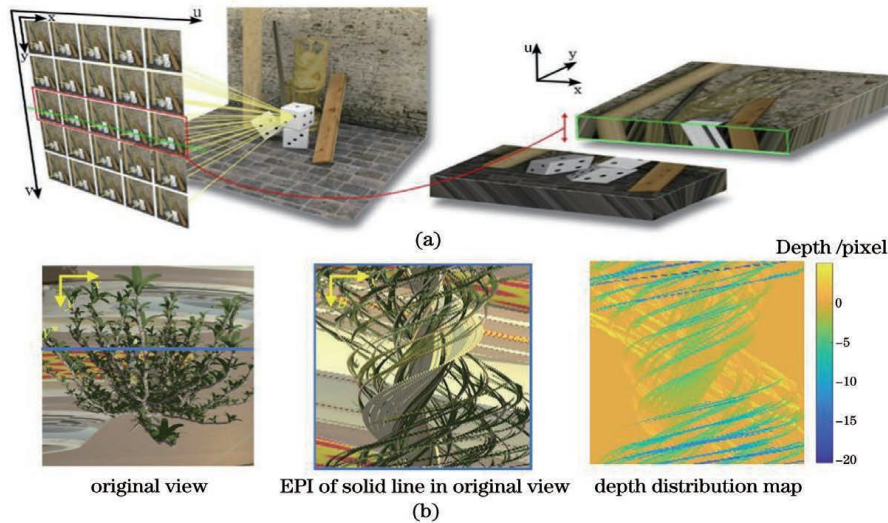


图 2 4D 光场的对极平面图表达。(a) 双平面光场的 EPI^[14]; (b) 环形光场的 EPI

Fig. 2 EPI representation of 4D light field. (a) EPI of two-parallel plane light field^[14]; (b) EPI of circular light field

式中: $d^{(x,y)}$ 为光场图像平面上点 (x, y) 处的深度值; B 为与相机内参相关的常数。在基于微透镜阵列和相机阵列的光场成像装置所采集到的光场中, 同一场景点在各个视点的深度是相同的, 即 $d^{(x,y)}$ 可视为常数。因此, 在双平面光场 EPI 中, 角度维和空间维有明显的线性相关特性, 这样的线性约束称为视差一致性。同时, 由于双平面光场 EPI 具有与深度相关的线性特性, 因此其傅里叶频谱能量分布受到最近和最远深度的限制, 而且对于场景中同一深度层的点(场景中所有深度为 Z_0 的点, 在 EPI 中显示为斜率相同的直线), 其频率支撑被限制在一条线上^[16]。环形光场的 EPI 也具有固有的几何特性。在极坐标下, 对于不同的旋转角度 θ , 世界坐标点在 xy 平面的投影如图 2(b) 所示^[51], 表达式为

$$\begin{cases} x = f \frac{R \sin(\theta + \phi)}{R_M - R \cos(\theta + \phi)} + x_c \\ y = f \frac{Y}{R_M - R \cos(\theta + \phi)} + y_c \end{cases}, \quad (5)$$

式中: $C(x_c, y_c)$ 为光心坐标; R_M 为相机旋转所在圆环的半径; ϕ 为初始相位; θ 为旋转角度, $\theta \in (0, 2\pi]$; R 为旋转半径。结合图 2(b) 和(4)式可知, 环形光场 EPI 的视差一致性表现为三维空间中的类正弦曲线, 利用 EPI 中的几何特性可简化诸多光场问题的分析。在双平面光场中, EPI 直线的斜率反映了场景点的视差信息, 通过提取 EPI 的直线特征能够直接对场景进行深度估计。Johannsen 等^[52] 利用 EPI 的几何特性构建了深度感知字典, 对 EPI 进行稀疏编码可以获得场景的深度信息。Wanner

等^[53] 利用结构张量估计了 EPI 的直线斜率, 进而估计场景深度。Zhou 等^[54] 认为不同空间尺度和方向的 EPI 块可能表现出不同的特征, 为此提出了一种尺度-方向感知的 EPI-Patch 网络模型并对 EPI 线性特性进行估计。Shin 等^[55] 利用 EPI 中的几何特性设计了一种端到端的神经网络结构, 同时对 EPI 在不同方向的空间和角度信息进行学习, 提取了 EPI 线的方向信息, 随后利用多层卷积网络对其进行融合优化, 获取了全局最优深度图。

另一方面, 由于 EPI 存在角度-空间一致性约束^[56], 利用其进行视点重建能够更好地保持光场角度一致性。因此, EPI 表达也广泛应用于光场角度超分辨。Wu 等^[57] 提出了基于深度卷积神经网络的光场重建方法, 利用 EPI 的几何特性将光场视点的重建问题转换为 EPI 角度维的信息恢复问题, 并利用已知的核函数对 EPI 进行模糊与去模糊, 可以实现大视差的光场重建。Guo 等^[50] 提出了一种基于渐进残差学习的高角度分辨率光场重建方法, 该方法将光场重建问题转化成稠密光场的 EPI 和稀疏光场的 EPI 之间残差的预测问题, 简化了光场重建问题的求解过程。Zhu 等^[58] 发现虽然 EPI 中包含两种不同类型的对极线, 即连续型和离散型, 但均可以使用统一的序列模型来描述, 因此提出了一种 CNN 与 RNN(Recurrent Neural Network) 相结合的深度学习网络结构, 使用该网络对 EPI 进行超分辨, 可以同时提高光场的角度和空间分辨率。该网络不需要视差, 直接对 EPI 上的对极线进行拟合与重建能够有效地对光场进行超分辨, 尤其是在

大视差区域,重建结果清晰锐利。Vagharshakyan 等^[59]结合了 EPI 频谱的特性对剪切波变换进行改进,获取了 EPI 频谱在剪切变换域中的稀疏表示,为此提出了一种基于自适应阈值的迭代正则化算法,采用该算法对 EPI 频谱的低频特征进行有效重构,可以达到光场重建的目的。此外,在光场显示方面,Xie 等^[60]根据 EPI 角度空间一致性线索,将真实场景和虚拟模型的光场自然地融合在一起,可以增强光场的 3D 显示。Kim 等^[61]通过分析 EPI 的几何特性,对光场三维显示中的畸变进行校正。

2.1.3 深度-尺度表达模型——聚焦栈表达

基于 EPI 的表达无法准确描述场景的纹理信息,而且直接从角度维-空间维表达光场仍然未能降低数据冗余。重聚焦图像不仅能够描述场景的纹理信息,其模糊特性也可以对物体的相对深度进行刻画。以场景深度分布为先验,仅需获取主要深度区间的重聚焦图像序列,即可对场景进行描述,降低光场数据的冗余^[62]。Ng^[3]从空域和频域的角度讨论了全透镜光圈拍摄的图像和 4D 光场之间的关系,将采集的光场重新投影到新的平面并进行积分,实现了数字重聚焦,表达式为

$$F(d, x, y) = \iint L[u, v, x + d(u - u_{ref}), y + d(v - v_{ref})] du dv, \quad (6)$$

式中: d 为聚焦深度所对应的视差; (u_{ref}, v_{ref}) 为重聚焦的参考视点。聚焦栈是光场在不同深度重聚焦^[62]后所形成的图像序列^[63]。以色列魏茨曼科学研究所的 Levin 等^[64]通过分析光场的先验信息,发现在无限孔径的情况下,4D 光场的能量主要集中在一个 3D 子集中,而聚焦栈的能量充分覆盖了该 3D 子集。因此,利用 3D 聚焦栈可对 4D 光场进行表达。如图 3 所示,其中点 Q 为空间点 X 在图像平面上的聚焦点,点 x_i 和 x_{i+1} 之间的距离为非聚焦平面上点 X 的模糊核直径,其随着聚焦

深度的变化而改变。不同聚焦深度图像上的模糊程度不同,而且对应于不同的尺度信息,因此聚焦栈是一种在深度-尺度空间对光场进行结构化分析与表达的手段。此外,Lin 等^[65]分析了聚焦栈在不同遮挡情况下的特性,如图 3 所示,其中 n 为位置编号, α 为期望的聚焦深度与原始聚焦深度的像距之比, F 为傅里叶变换算, F' 为聚焦深度。在朗伯的假设下,即使存在噪声和采样不足,聚焦栈关于聚焦深度具有对称性,因此 3D 聚焦栈序列具有稀疏性。

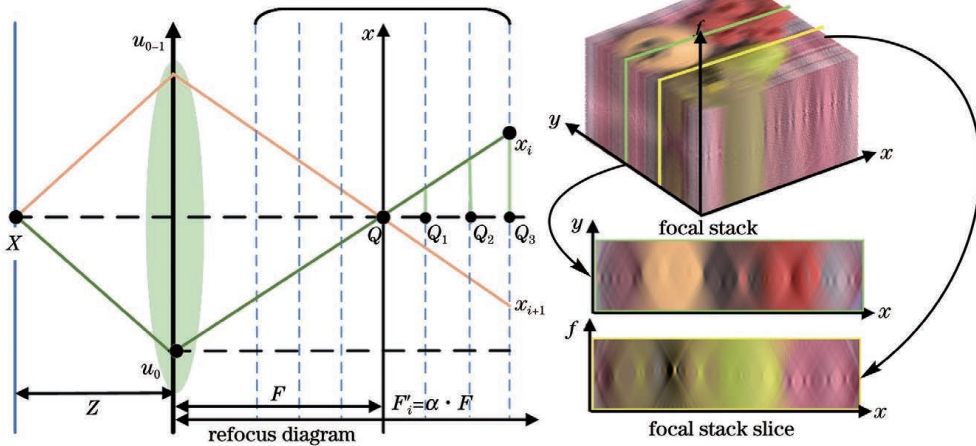


图 3 4D 光场的聚焦栈表达

Fig. 3 Focal stack representation of 4D light field

聚焦栈表达同样广泛应用于光场重建、深度估计、光场显示、显微成像等相关应用。在光场重建方面,Levin 等^[64]提出利用三维聚焦栈对 4D 光场进行频谱补充的方法,对稀疏光场进行角度域的稠密重建。Sakamoto 等^[19]将聚焦栈看作光场降维的中间表示,并采用基于小波的方法对聚焦栈进行编码,

然后采用 Levin 等^[64]描述的线性视图合成方法从聚焦栈中重建光场。Huang 等^[66]以聚焦栈作为输入,通过三个卷积神经网络对光场进行重建,第一个网络估计全聚焦图像,第二个网络从聚焦栈和全聚焦图像中估计场景深度,第三个网络完成最终的渲染重建任务。Liu 等^[67]在投影建模的基础上,采用

层析成像方法^[68]推导了光场重建问题的滤波反投影法^[69]和 Landweber 迭代方法^[70],实现了光场重建。利用聚焦栈进行深度估计,其本质是通过分析聚焦栈中聚焦或离焦度来推导场景深度^[71-75],但是受限于聚焦和离焦法本身的计算精度,这些算法并不能在遮挡边缘等区域实现性能的突破。为了克服该缺点,Strecke 等^[76]选用非遮挡视点构建聚焦栈来进行深度估计。另外,Rajagopalan 等^[77]通过计算点扩展函数的一致性来进行视差估计,但是该方法在大模糊核以及大视差的情况下效果不佳。Lin 等^[65]使用了聚焦栈对称性计算场景深度。聚焦栈表达还可以用来进行光场显示,Takahashi 等^[78]以聚焦栈作为输入,设计了一种新型张量光场显示器,该显示器仅由位于背光前面的几个光衰减层组成,可以同时从多个方向(角度)观看,在保证输出光场质量的同时大大降低了获取数据的成本。光场特有的角度维采样能够在不同聚焦深度成像,为实时显微三维成像提供了数据保证,有效提升活体细胞的观测能力^[79-82]。

2.1.4 深度-纹理表达模型——光场超像素、图/超图、多平面图像表达

超像素是计算机视觉的一个重要工具,其主要作用在于将图像中颜色相似和距离相近的同一目标

中的像素聚成有意义的图像块,以提高后续算法的精度与速度^[83]。传统的超像素表达将光场图像看作一个二维图像序列,忽略了 4D 光场中视点之间的关联性。朱昊^[56]在光场空间角度一致性理论的指导下,将光场超像素定义为 3D 空间中一组颜色相似和深度相近的连通区域所发出的光线集合,由此将光场超像素表达为一种深度-纹理表达模型。设 R 为三维空间中一块颜色相近的连通物体表面,4D 光场为 $L(u, v, x, y)$,则 R 区域的光场超像素 $S_R(u, v, x, y)$ 为

$$s_R(u, v, x, y) = \bigcup_{i=1}^{|\mathbf{R}|} L(u_{P_i}, v_{P_i}, x_{P_i}, y_{P_i}), \quad (7)$$

式中: $L(u_{P_i}, v_{P_i}, x_{P_i}, y_{P_i})$ 为从三维空间 R 区域的第 i 个点 P_i 发出的光线集合,满足 $L(u_{P_i}, v_{P_i}, x_{P_i}, y_{P_i}) \subseteq L_F(u, v, x, y)$, $|\cdot|$ 为集合中的元素数目, L_F 为原四维光场。

图 4 为光场成像与传统成像在离焦和遮挡下的区别^[56]。在传统的薄透镜成像模型下, p 点的像过于模糊而且难以进行分割。在光场成像中,每一根从不同角度发出的光线都被清晰记录,因此光场超像素能够从根本上消除遮挡和离焦情况中出现的分割歧义性问题。

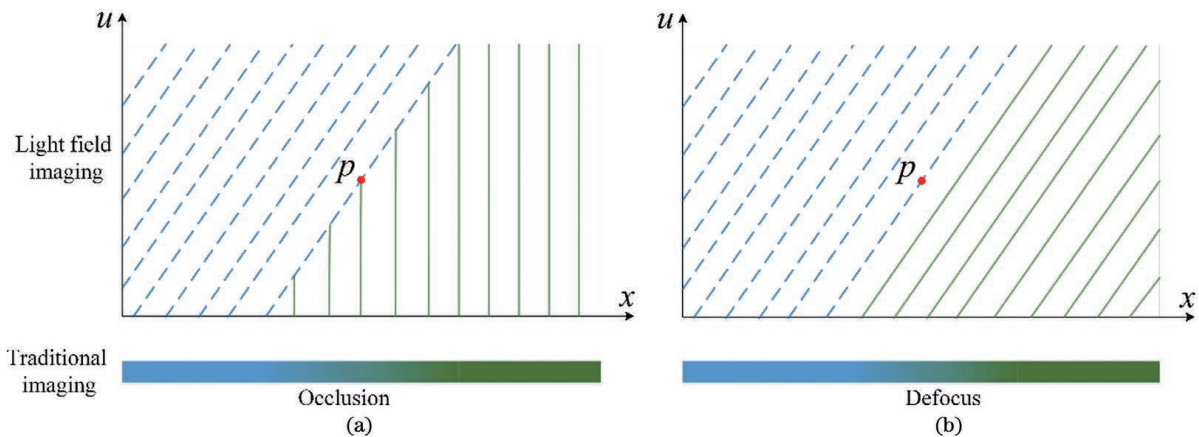


图 4 光场成像与传统成像在不同情况下的对比^[56]。(a)遮挡;(b)离焦

Fig. 4 Comparison between light field imaging and traditional imaging in different situations^[56]. (a) Occlusion; (b) defocus

同时,根据光场超像素的定义,超像素中的每条光线应当具有“重聚焦不变性”,即在重聚焦前后,每条光线所在的超像素编号应当保持不变,这是因为在重聚焦前后,每根光线所描述的三维空间点的位置与纹理没有发生变化,重聚焦并不会改变光场中所记录的信息。光场超像素还具有“切片完整性”,即在光场超像素中,空间中同一区域在不同视点下的成像属于同一个超像素,如图 5 所示^[84]。

光场超像素表达主要用于光场编辑及场景流估计等应用。Zhu 等^[84]利用了中心视点的视差图并通过级联方式将中心视点与边缘视点的分割结果连接起来,为此提出了一个光场超像素分割方法并给出了光场超像素自相似性的评价指标。该方法在计算光场像素间距离的过程中,将视差因素考虑在内,因此不仅具有良好的边缘贴合性,还具有“重聚焦不变性”及“切片完整性”。Khan

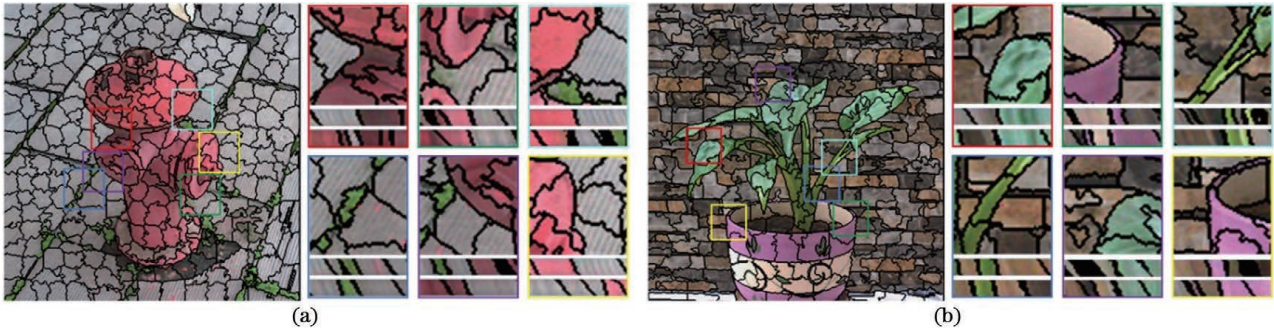


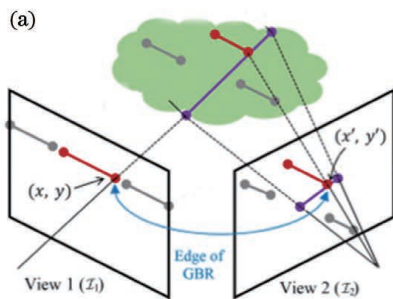
图 5 4D 光场的超像素表达及其局部放大图^[84]。(a) 图像 1; (b) 图像 2

Fig. 5 Superpixel representation of 4D light field and its local magnification^[84]. (a) Image 1; (b) image 2

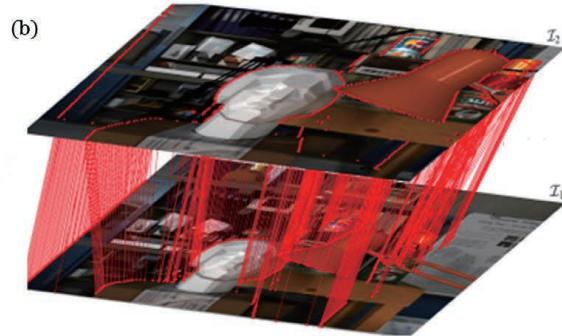
等^[85]通过分割 EPI 隐式地计算光场深度,然后聚类得到光场超像素分割结果,该算法能够保持分割过程中的视差一致性。光场超像素表达还被应用于场景流估计,Zhu 等^[86]对单个光场中的超像素在 3D 空间进行斜面建模,发现每个视点下光场超像素切片的运动能够有效映射到中心视点,通过计算相邻光场间中心视点的光流,能够实现光场全视点的场景流估计。在深度估计方面,Chen 等^[87]通过超像素正则化来检测部分遮挡边界区域,从而对光场进行深度估计,该方法能够准确地

估计遮挡边缘的深度。

基于图的光场表达模型将场景几何信息转换为视点间的像素关联。图的顶点表示颜色信息,边表示深度(视差)信息,用来连接不同视点的相应像素^[88]。图表达不仅包含了光场的深度-纹理信息,更是一种紧凑的光场表达方式。图 6 为光场图表达的示意图和真实数据上的示例^[88],顶点对应于视图中的像素 (x, y) ,边连接视图 1 中的像素 (x, y) 与其在另一视图下的投影 (x', y') ,示例图中视点图之间的连线为图表达中的边。



Schematic



An example of the graph for real dataset.

图 6 4D 光场图的表达^[88]。(a) 光场图的表达; (b) 真实数据

Fig. 6 Representation of 4D light field graph^[88]. (a) Expression of light field pattern; (b) real data

光场超图表达是图表达的一般化^[89],本质是语义层面上对光场更简洁的表达方法。超图中的一条边可以连接多个顶点,因此光场超图表达的一般形式为 $H = (V, E)$ ^[90], 其中

$$\begin{cases} V = \left\{ \bigcup_{i, (u, v)} r_i^{(u, v)} \right\} \\ E = \left\{ \left\{ \bigcup_{(u, v)} r_i^{(u, v)}, p_i \right\}, \left\{ \bigcup_{i, j, (u, v)} r_i^{(u, v)}, r_j^{(u, v)} \right\} \right\} \end{cases}, \quad (8)$$

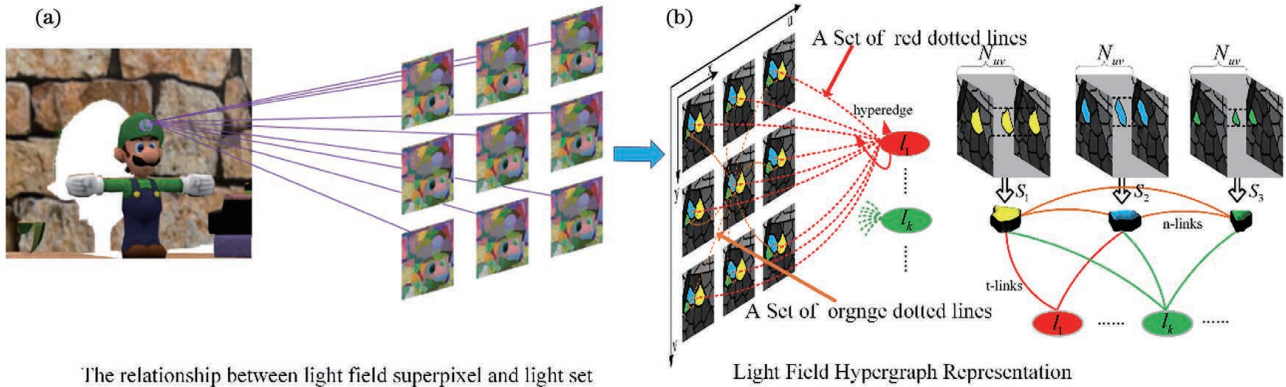
式中: V 为超图的顶点集合; E 为光场超图边的集合; $r_i^{(u, v)}$ 为在 (u, v) 视点捕获的第 i 个场景点发出的光线; j 为 i 的相邻顶点。 $\left\{ \bigcup_{i, (u, v)} r_i^{(u, v)} \right\}$ 为光场超

图边集合中所包含的光线对应关系,即多条光线与唯一场景点的对应关系; 光线集的邻接关系 $\left\{ \bigcup_{i, j, (u, v)} r_i^{(u, v)}, r_j^{(u, v)} \right\}$ 即为场景点相对关系在光线集上的表现。

上述光场超图表达中,顶点集合中的元素为单条光线,这会导致光场超图的复杂度非常高,而且精确地匹配不同视点中的相同光线束十分具有挑战性。为了解决这些问题,Lü 等^[91]利用光场超像素对光场超图表达进行简化。根据(7)式对超像素的定义,将超像素作为超图顶点集的子集,同时处理场景中一块区域,而不是仅仅处理一个场景点的信

息。接下来,根据超像素的“重聚焦不变性”^[84]以及光场超像素与光线集的关系来合并相似顶点,进一步对光场超图表达进行了简化。4D 光场的超图表达如图 7 所示^[91],其中红色虚线形成了一个 t -连接

(终端连接),橙色虚线构成一个 n -连接(邻域连接)^[92-93], S_i 为不同的超像素块,标签 l_k 为进一步简化后的分类标签,其中 t 和 n 为终端连接与邻域连接个数, k 为标签类别数。



The relationship between light field superpixel and light set

Light Field Hypergraph Representation

图 7 4D 光场的超图表达^[91]。(a)光场超像素与光集的关系;(b)光域超图表示

Fig. 7 Hypergraph representation of 4D light field^[91]. (a) Relationship between light field superpixel and light set; (b) light field hypergraph representation

光场的图和超图表达主要用于光场分割、重构及压缩等应用。Maughey 等^[94]使用密集视差来求解图的顶点连接关系以构建图表达模型,使用该模型将水平排布的多视点图像进行压缩与重构。在此基础上,Su 等^[88]考虑了更一般化相机配置下的图表达,该表达进一步描述了场景的几何信息,具有更高的重构质量和更低的编码率。Su 等^[95]通过额外增加视点残差信息提高了基于图表达模型的光场重构质量。光场图表达模型的复杂性会导致光场压缩与重构效率较低,在结合光场超像素^[84]信息以后,Rizkallah 等^[96]将超像素作为局部图变换的基础进行数据压缩,并且在重构框架中进行了视点合成的测试。最近,Lü 等^[91]在图表达和超像素表示的基础上,从光线与光线集的角度刻画了光场并得到了光场的超图表达模型,其表达方式及简化过程能够为多种光场处理任务提供理论解释,如光场超像素分割和光场压缩等。

近年来,随着深度学习技术的发展,Zhou 等^[97]提出了一种基于学习的光场多平面图像(Multiplane Images, MPI)表达,如图 8 所示。MPI 表达是从两张具有已知相机参数的图像(使用图 8 的两台相机姿势拍摄而得),利用深度神经网络来推理场景的隐式深度表达。该表达由多层图像构成,每一层图像对应于某一确定深度 d ,该层图像由一个 RGB (Red, Blue, Green) 彩色图像 C_d 和一个透明度 α_d 构成,这些层的集合就是光场 MPI 表达,即 (C_o, α_o) ($o=1, 2, \dots, d, \dots, D$),其中 D 为深度平面的数量。

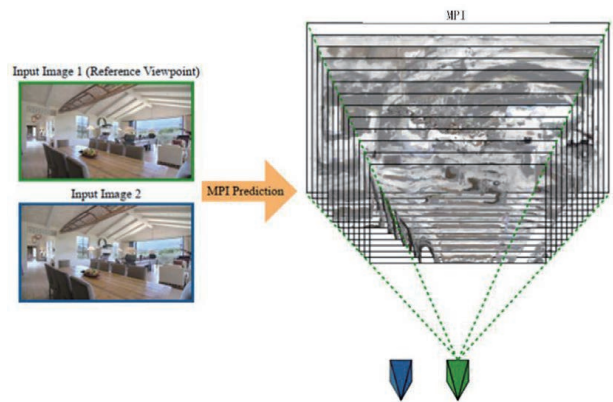


图 8 光场的多平面图像表达^[98]

Fig. 8 Multiplane image representation of light field^[98]

由于 MPI 同时包含了场景的深度和纹理信息,因此 MPI 也是一种深度-纹理表达模型。

随后,Srinivasan 等^[98]在 MPI 空域表达的基础上,分析了 MPI 的频率特性,图 9(a)为固定 MPI 的 y 维,只有 x 维和 z 维(深度)变化的二维切片示意图,其中 v_o 为参考视点,图 9(b)为对应的频域示意图。在空域中 MPI 通过投影融合可以得到新视点,表现在频谱中则为对应二维频谱的一维切片,而且离参考视点越远所对应切片的倾斜度越大。

目前,光场 MPI 表达主要用于视点重建等相关应用^[97-102]。Zhou 等^[97]通过 MPI 的后向投影和透明度 α 融合可以得到新视点,在视点内插的同时还能实现视点外扩。Srinivasan 等^[98]使用 3D 卷积神经网络提高了视点外扩的能力。同时,2D 流向量的使用使其在遮挡区域也能够获得较为精确的重建结

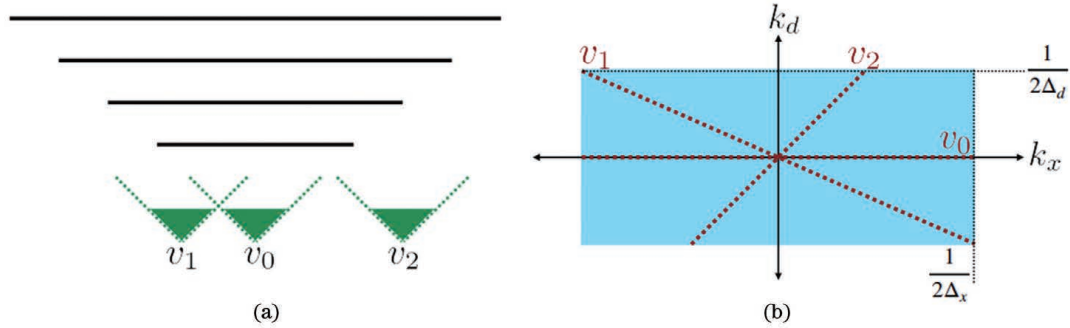


图 9 2D MPI 及其频域表达示意图^[98]。(a)空间域的 MPI;(b)频域的 MPI

Fig. 9 Schematic of 2D MPI and its frequency domain representation^[98]. (a) MPI in space domain; (b) MPI in frequency domain

果。Mildenhall 等^[99]通过多个 MPI 信息的融合进一步提升了光场重建的性能。从 MPI 的构建出发, Flynn 等^[100]基于学习的梯度下降算法从一组稀疏的相机视点生成 MPI 并进行光场重建,该算法中相关梯度信息编码所具有的可见性使网络能够对遮挡区域进行有效推导,提高了算法在物体边界、光线反射、毛细结构和高深度复杂性等极具挑战区域中纹理重建的能力。Li 等^[101]和 Tucker 等^[102]分别提出了利用单张图像构建 MPI 并进行视点合成的方法。

2.2 频域光场的表达模型

2.2.1 光场傅里叶切片表达

在传统的傅里叶切片理论的基础上^[103-104]结合光场频域特性,Ng^[105]提出了光场傅里叶切片模型,给出了 4D 光场的 2D 傅里叶切片表示。在频域中,聚焦在不同深度的图像等价于 4D 光场频谱上不同角度的切片。通过 4D 光场切片的处理,可以实现快速的数字重聚焦。傅里叶切片定理如图 10 所

示^[105]。

传统的 2D 傅里叶切片理论是 CT(Computed Tomography)成像的理论基础^[106],空域中的一张脑部 CT 图像,通过积分投影再一维傅里叶变化得到的频域数据等价于该 CT 图像经过 2D 傅里叶变化后的一维切片^[107],如图 10(a)所示。图 10(b)将一维切片扩展到 N 维^[108],分析了 N 维函数 G_N 和其 M 维积分投影 G_M 以及其傅里叶谱 B_N 到 B_M 的变化关系。广义傅里叶切片的数学表达式为

$$F^M \circ I_M^N \circ \mathbf{B} \equiv S_M^N \circ \frac{\mathbf{B}^{-T}}{|\mathbf{B}^{-T}|} \circ F^N, \quad (9)$$

式中: I_M^N 为将 N 维函数降为投影到 M 维的积分投影算子; S_M^N 为将 N 维函数剪切到 M 维的剪切算子,通过多余项置零来完成; \mathbf{B} 为变换矩阵,指导具体的积分投影和剪切操作过程; \mathbf{B}^{-T} 为 \mathbf{B} 的逆转置; $|\mathbf{B}|$ 为 \mathbf{B} 的行列式; \circ 为复合运算符; F 为傅里叶变换。

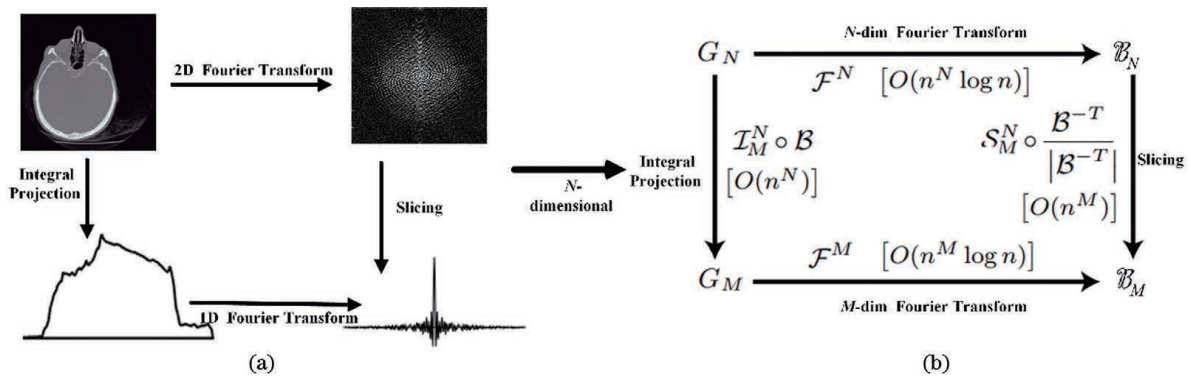


图 10 傅里叶切片定理^[105]。(a)传统二维傅里叶切片定理;(b)广义傅里叶切片定理

Fig. 10 Fourier slice theorem^[105]. (a) Traditional 2D Fourier slice theorem; (b) generalized Fourier slice theorem

Ng^[105]将广义傅里叶切片定理应用于光场成像,得到了基于傅里叶理论的光场成像定理,即 4D 频域中傅里叶光场的二维切片操作等价于空域中 4D 光场的积分操作,如图 11 所示。该过程的数学

定义为

$$\frac{1}{\alpha^2 F^2} I_2^4 \circ \mathbf{B}_{ar}[L_F] \equiv \frac{1}{F^2} F^{-2} \circ S_2^4 \circ \mathbf{B}_{ar}^{-T} \circ F^4, \quad (10)$$

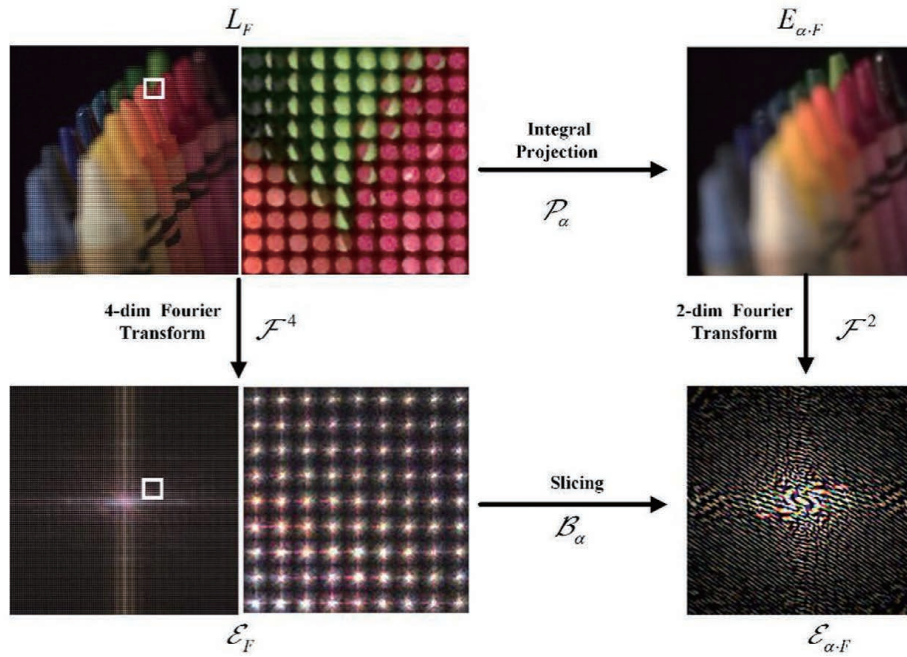


图 11 傅里叶切片成像定理^[105]

Fig. 11 Fourier slice photography theorem^[105]

式中： α 为期望的聚焦深度与原始聚焦深度的像距之比，用来确定积分投影的深度平面，如图 3 所示； F^2 为 2D 傅里叶变换； F^{-2} 为 2D 傅里叶逆变换； B_α 为变换矩阵，则

$$B_{ar} = \begin{bmatrix} \alpha & 0 & 1 - \alpha & 0 \\ 0 & \alpha & 0 & 1 - \alpha \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (11)$$

$$B_{ar}^{-1} = \begin{bmatrix} \frac{1}{\alpha} & 0 & 1 - \frac{1}{\alpha} & 0 \\ 0 & \frac{1}{\alpha} & 0 & 1 - \frac{1}{\alpha} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (12)$$

理论上，在频域中通过多次切片操作可以得到包含场景所有深度的二维切片组。在无限孔径的假设下，4D 光场的能量集中于该二维切片组，因此可以采用傅里叶切片来表达 4D 光场。

重聚焦操作的本质是光场在聚焦深度上的积分投影，因此光场傅里叶切片表达可以被用于光场数字的重聚焦。Ng^[105] 在频域中对 4D 光场进行不同方向的二维切片，经过傅里叶逆变换可以得到不同深度的重聚焦图像。由于减少了积分操作，重聚焦效率得到了较大的提升。赵青青等^[109] 提出了一种基于傅里叶切片的高分辨率数字对焦方法，通过提取子孔径图像来组成高分辨率的子孔径图像阵列，

再对该子孔径图像阵列进行傅里叶切片以得到重聚焦图像。Levin 等^[110] 在光场傅里叶切片的基础上分析光场频谱能量的主要分布，设计了一套可用于扩展景深的光场成像系统。此外，Shi 等^[111] 和 Levin 等^[64] 通过分析傅里叶切片的稀疏性，分别实现了更小样本的光场重建。根据傅里叶切片和重聚焦图像的关系，还可以实现显微场景的快速重聚焦和 3D 重建^[112]。

2.2.2 傅里叶视差层表达

以傅里叶切片定理为基础，Le Pendu 等^[113] 提出了傅里叶视差层 (Fourier disparity layer, FDL) 表达，并使用其来处理 and 渲染光场。图 12(a) 为 FDL 的示意图，FDL 依据视差 (深度) 将场景划分成一组离散的视差层 (l) 以实现光场采样，视差层可以由不同类型的光场输入构建，如一组子孔径图像和一组聚焦在不同深度的图像或者两者的组合。同时，通过梯度下降方式可以估计重建光场所需的最优视差层数^[113]。图 12(b) 为真实数据上 FDL 每一层的表示，右下角方框为对应的功率谱。在视觉上 FDL 和聚焦栈表达的区别：FDL 表达焦点外的区域趋于消失，而焦点内的区域具有更强的对比度^[113]。

不失一般性，为了简化符号，在包含一个空间维 x 和一个角度维 u 的二维光场 $L(x, u)$ 下的 FDL 构造过程如下。

在朗伯且无遮挡的先验下，场景可以被分成 n 个空间区域，每个区域 Ω_k 对应一个固定的视差

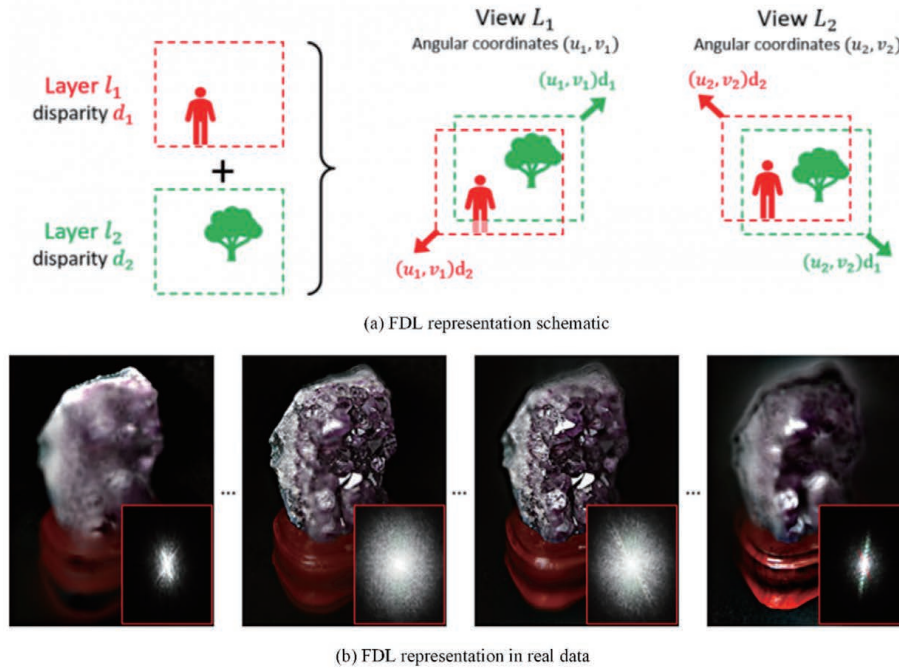


图 12 傅里叶视差层表达^[113]。(a) FDL 表示原理图；(b) 真实数据中的 FDL 表示

Fig. 12 Fourier parallax layer representation^[113]. (a) FDL representation schematic; (b) FDL representation in real data

值 d_k , 表达式为

$$\forall k \in [1, n], \forall (x, u) \in \Omega_k \times \mathbb{R},$$

$$L(x - ud_k, u) = L(x, 0). \quad (13)$$

如(12)式所示,在参考视点 $u=0$ 下,对于 Ω_k ,根据其视差 d_k ,可被投影到任意其他视点,且投影前后的纹理强度不变。在此基础上,Le Pendu 等^[113]通过推导得到光场傅里叶变换可分解为

$$\hat{L}(\omega_x, \omega_u) = \sum_k \delta(\omega_u - d_k \omega_x) \hat{L}^k(\omega_x), \quad (14)$$

式中: (ω_x, ω_u) 为二维光场的频域坐标; δ 为狄利克雷函数; \hat{L}^k 为

$$\hat{L}^k(\omega_x) = \int_{\Omega_k} \exp(-2i\pi x \omega_x) L(x, 0) dx. \quad (15)$$

每一层 \hat{L}^k 为中心视图中只考虑视差 d_k 的 Ω_k 区域的傅里叶变换。因此(13)~(15)式被称为光场的 FDL 表达。

FDL 表达主要用于光场重建和压缩等应用。Le Pendu 等^[113]通过 FDL 中视差层的平移加和操作,可以得到光场任意视点 (u, v) 下的图像重建结果,如图 12(a)所示。随后,Le Pendu 等^[114]进一步发现,如果空间采样具有规则性,那么可以通过线性优化来构建 FDL 表达,因此可以通过求解傅里叶域中简单的线性方程对光场进行重建。Dib 等^[115]首先对光场子集进行编码,用来构建 FDL 表达模型,从中预测第二个视图子集;然后对预测的视图子

集通过残差进行编解码,可以获得更大的视图集合,并利用此集合进一步细化 FDL 表达,从而得到更精确的视点预测;最后重复上述过程,直到完整的光场视图被压缩编码。Le Pendu 等^[116]提出了一种存储 FDL 的树形结构,该结构的传输量远小于原始 4D 光场。

2.2.3 对极聚焦谱表达

同样以傅里叶切片定理为基础,Li 等^[117]根据视差线索 $f \in [d_{\min}, d_{\max}]$ 对 EPI 频谱进行切片,然后按照视差变化进行重排列,最后在 f 轴上进行一维傅里叶变换,从而得到了光场的对极聚焦谱(Epipolar Focus Spectrum, EFS)表达 $[F_{\text{EFS}}(\omega_f, \omega_x)]$ 如图 13 第二行所示。不失一般性,二维光场 $L(x, u)$ 下的 EFS 数学表达构造过程为

$$\hat{F}(f, \omega_x) = \hat{L}(-f\omega_u, \omega_x), \quad (16)$$

$$F_{\text{EFS}}(\omega_f, \omega_x) = F_{T,f}[\hat{F}(f, \omega_x)], \quad (17)$$

式中: \hat{F} 为 EPI 频谱沿视差 f 的切片; \hat{L} 为 2D 光场频谱; $F_{T,f}$ 为沿着 f 轴的一维傅里叶变换。根据视差 f 对 2D 光场频谱 $\hat{L}(\omega_x, \omega_u)$ 进行切片,可以得到 $\hat{F}(f, \omega_x)$ 。同时,从图 13 第一行出发,还可以从聚焦栈构造 EFS,构造过程为

$$F(f, x) = \int L[u, x + f(u - u_{\text{ref}})] du, \quad (18)$$

$$F_{\text{EFS}}(\omega_f, \omega_x) = F_T[F(f, x)], \quad (19)$$

式中: $F_T(\cdot)$ 为二维傅里叶变换; $F(f, x)$ 为聚焦在

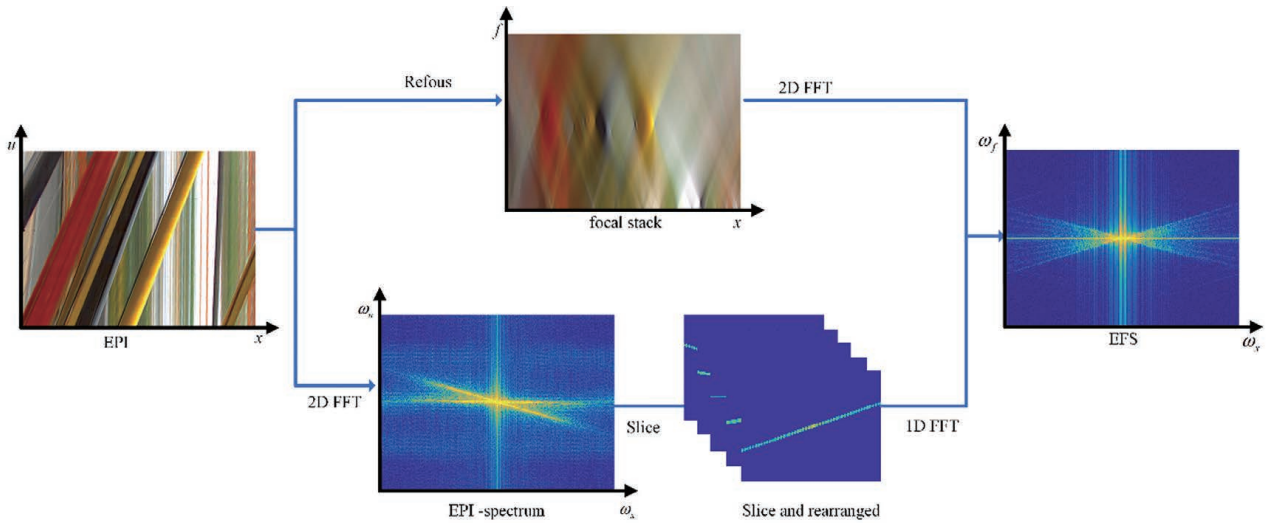


图 13 对极聚焦谱 EFS 的两种构造方式^[117]

Fig. 13 Two construction methods of polar focus spectrum EFS^[117]

不同深度的重聚焦图像所组成的聚焦栈。

Li 等^[117]通过分析发现, EFS 的形状仅与重聚焦参数和视点个数有关, 即 EFS 中每一条频谱线对应于光场的一个视点。根据该特性, 通过对 EFS 进行频谱重建可以实现光场的抗混叠绘制以及视点重

建等任务。

2.3 总结

综合以上分析, 不同表达模型描述了光场的不同特性并应用于不同任务。表 1 为上述表达模型各项特性的对比分析, 包括光场的相关应用。

表 1 不同模型的各项特性和应用

Table 1 Characteristics and applications of different models

Model	Depth dependence	Working domain	Occlusion robustness	Complecity	Redundancy	Light field application		
						Depth recovery	Reconstruction	Others
Sub-aperture image	No	Spatial-angular	Low	Low	High	[28-31]	[33-47]	Microscopic imaging ^[48] 3D measurement ^[49]
EPI	No	Spatial-angular	Medium	Low	High	[52-55]	[56-59]	3D display ^[60-61] Display ^[78]
Focal stack	No	Spatial-scale	Medium	Medium	Medium	[65,71-77]	[17,64,66-69]	Microscopic imaging ^[79-82]
Superpixel	Explicit depth	Spatial-angular-depth	High	Medium	Medium	[87]	-	Editing ^[84-86]
Graph/hypergraph	Explicit depth	Spatial-angular-depth	High	High	Low	-	[88,94-95]	Editing ^[91] Compression ^[96]
MPI	Implicit depth	Spatial-depth	High	High	Medium	-	[97-102]	-
Fourier slice	No	Frequency	Medium	Low	High	-	[64,111]	Refocus ^[105,109] Depth of field extension ^[110] Microscopic imaging ^[112]
FDL	No	Frequency-depth	Medium	Low	Low	-	[113-114]	Compression ^[115-116]
EFS	No	Frequency-focal	Medium	Low	Medium	-	[117]	Refocus ^[117]

3 结束语

光场成像能够突破传统成像的局限,具有变视点成像、层析成像以及全聚焦成像等诸多优异的成像性能,已成为诸多学科突破现有观测瓶颈的重要手段之一。原始的光场数据因具有高维结构性和数据冗余性的缺点而限制了其应用。从光场数据的深度、纹理、几何、尺度及频谱等特性出发,构建不同的光场表达模型,能对光场进行更简洁且高效的表达。

通过对现有方法的梳理发现现有的光场表达模型仅聚焦于光场的某一个或者某几个特性,无法完备地表达光场各个维度的信息。如何针对更高维度的数据寻求简洁且高效的光场全特性表达模型是未来光场理论研究的重点之一。此外,现有的光场表达模型忽略了时间维度,因此仅适用于静态光场,缺乏对高速时变场景中动态变化光场信息的刻画与建模能力,限制了光场成像理论及技术的应用领域。为了延拓光场成像的应用范围,需在现有深度、纹理、几何、尺度及频谱等信息外,从动态光场的采集成像出发,进一步分析时变光场不同时刻非线性变换光线间的关联,构建对光场连续非线性变换的高效编码表示理论及相应的可压缩先验,为高质量动态光场的重构与分析提供理论基础。

综上,光场表达模型的提炼与深化是解决诸多计算机视觉问题的基石,以此为基础的诸类算法已显著提升显微观测、流体分析、天体检测及影视制作等领域的精度,有效推动了生物、物理及天文等相关学科的发展。对于不同光场表达的推导、特性以及相关应用的总结分析具有重要的理论意义和应用价值。

参 考 文 献

- [1] Suo J L, Liu Y B, Ji X Y, et al. Computational photography: keys, methods and applications [J]. *Acta Automatica Sinica*, 2015, 41(4): 669-685.
索津莉, 刘焯斌, 季向阳, 等. 计算摄像学: 核心、方法与应用[J]. *自动化学报*, 2015, 41(4): 669-685.
- [2] Levoy M. Light fields and computational imaging [J]. *Computer*, 2006, 39(8): 46-55.
- [3] Ng R. Digital light field photography[D]. Stanford: Stanford University, 2006.
- [4] Sugita K, Takahashi K, Naemura T, et al. Focus measurement on programmable graphics hardware for all in-focus rendering from light fields [C] // *IEEE Virtual Reality 2004*, March 27-31, 2004, Chicago, IL, USA. New York: IEEE Press, 2004: 255-256.
- [5] Kalantari N K, Ramamoorthi R. Deep high dynamic range imaging of dynamic scenes[J]. *ACM Transactions on Graphics*, 2017, 36(4): 11487-11505.
- [6] Dansereau D G, Pizarro O, Williams S B. Linear volumetric focus for light field cameras [J]. *ACM Transactions on Graphics*, 2015, 34(2): 15.
- [7] Viganò N, der Sarkissian H, Herzog C, et al. Tomographic approach for the quantitative scene reconstruction from light field images [J]. *Optics Express*, 2018, 26(18): 22574-22602.
- [8] Fattal D, Peng Z, Tran T, et al. A multi-directional backlight for a wide-angle, glasses-free three-dimensional display [J]. *Nature*, 2013, 495(7441): 348-351.
- [9] Bouman K L, Johnson M D, Zoran D, et al. Computational imaging for VLBI image reconstruction [C] // *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 913-922.
- [10] Raskar R, Tumblin J. Computational photography [C] // *ACM SIGGRAPH 2007 courses on SIGGRAPH '07*, August 5-9, 2007, San Diego, California. New York: ACM Press, 2007.
- [11] 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.
- [12] 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.
- [13] Cai Z W, Liu X L, Pedrini G, et al. Accurate depth estimation in structured light fields [J]. *Optics Express*, 2019, 27(9): 13532-13546.
- [14] 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.
- [15] Zhu H, Wang Q, Yu J Y. Light field imaging: models, calibrations, reconstructions, and applications [J]. *Frontiers of Information Technology & Electronic Engineering*, 2017, 18(9): 1236-1249.
- [16] Momoni Y, Yamamoto K, Yokote Y, et al. Light field mirage using multiple flat-panel light field displays[J]. *Optics Express*, 2021, 29(7): 10406-10423.
- [17] Park J H, Lee S K, Jo N Y, et al. Light ray field capture using focal plane sweeping and its optical reconstruction using 3D displays[J]. *Optics Express*, 2014, 22(21): 25444-25454.

- [18] 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 Press, 2000: 307-318.
- [19] Sakamoto T, Kodama K, Hamamoto T. A study on efficient compression of multi-focus images for denselight-field reconstruction [C] // 2012 Visual Communications and Image Processing, November 27-30, 2012, San Diego, CA, USA. New York: IEEE Press, 2012: 13249324.
- [20] Adelson E H, Bergen J R. The plenoptic function and the elements of early vision [M] // Landy M, Movshon J K. Computational models of visual processing. Cambridge: MIT Press, 1991.
- [21] 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.
- [22] Birklbauer C, Bimber O. Panorama light-field imaging [J]. Computer Graphics Forum, 2014, 33(2): 43-52.
- [23] 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 York, NY, United States. New York: ACM Press, 1996: 31-42.
- [24] Zhou Z L. Research on light field imaging technology [D]. Hefei: University of Science and Technology of China, 2012.
周志良. 光场成像技术研究 [D]. 合肥: 中国科学技术大学, 2012.
- [25] Gortler S J, Grzeszczuk R, Szeliski R, et al. The lumigraph [C] // Proceedings of the 23rd annual conference on Computer graphics and interactive techniques-SIGGRAPH'96, August 4-9, 1996, New York, NY, United States. New York: ACM Press, 1996: 43-54.
- [26] Zhang Q, Ling J B, Wang Q, et al. Ray-space projection model for light field camera [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 10113-10121.
- [27] Zhang Q, Wang Q, Li H D, et al. Ray-space epipolar geometry for light field cameras [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 5949: 1-14.
- [28] 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, December 1-8, 2013, Sydney, NSW, Australia. New York: IEEE Press, 2013: 2792-2799.
- [29] 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), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 1547-1555.
- [30] Wang T C, Efros A A, Ramamoorthi R. Occlusion-aware depth estimation using light-field cameras [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 3487-3495.
- [31] Zhu H, Wang Q, Yu J Y. Occlusion-model guided antiocclusion depth estimation in light field [J]. IEEE Journal of Selected Topics in Signal Processing, 2017, 11(7): 965-978.
- [32] Shum H Y, Chan S C, Kang S B. Image-based rendering [M]. Boston: Springer, 2008.
- [33] Eisemann M, de Decker B, Magnor M, et al. Floating textures [J]. Computer Graphics Forum, 2008, 27(2): 409-418.
- [34] Goesele M, Ackermann J, Fuhrmann S, et al. Ambient point clouds for view interpolation [C] // ACM SIGGRAPH 2010 papers on-SIGGRAPH'10, July 26-30, 2010, Los Angeles, California. New York: ACM Press, 2010: 95.
- [35] Chaurasia G, Sorkine O, Drettakis G. Silhouette-aware warping for image-based rendering [J]. Computer Graphics Forum, 2011, 30(4): 1223-1232.
- [36] 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.
- [37] Penner E, Zhang L. Soft 3D reconstruction for view synthesis [J]. ACM Transactions on Graphics, 2017, 36(6): 235.
- [38] Srinivasan P P, Wang T Z, Sreelal A, et al. Learning to synthesize a 4D RGBD light field from a single image [C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2262-2270.
- [39] Wang T C, Zhu J Y, Kalantari N K, et al. Light field video capture using a learning-based hybrid imaging system [J]. ACM Transactions on Graphics, 2017, 36(4): 133.
- [40] Bishop T E, Favaro P. The light field camera:

- extended depth of field, aliasing, and superresolution [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(5): 972-986.
- [41] Wu J. Introduction to convolutional neural networks [R]. Nanjing: Nanjing University, 2017, 5: 23.
- [42] Wang Y L, Liu F, Wang Z L, et al. End-to-end view synthesis for light field imaging with pseudo 4DCNN [M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11206: 340-355.
- [43] 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.
- [44] Flynn J, Neulander I, Philbin J, et al. Deep stereo: learning to predict new views from the world's imagery [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 5515-5524.
- [45] 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.
- [46] 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 Press, 2015: 57-65.
- [47] Zhao Y Y, Shi S X. Light-field image super-resolution based on multi-scale feature fusion [J]. *Opto-Electronic Engineering*, 2020, 47(12): 56-66.
赵圆圆, 施圣贤. 融合多尺度特征的光场图像超分辨率方法 [J]. *光电工程*, 2020, 47(12): 56-66.
- [48] Yang F F, Li H, Peng J, et al. Research on microscopic imaging of high resolution light field based on graph regularization [J]. *Acta Optica Sinica*, 2021, 41(9): 0918001.
杨飞凡, 李晖, 彭晶, 等. 基于图正则化的高分辨率光场显微成像研究 [J]. *光学学报*, 2021, 41(9): 0918001.
- [49] Wu J L, Guo Z H, Chen X F, et al. Three-dimensional measurement method of light field imaging based on deep learning [J]. *Chinese Journal of Lasers*, 2020, 47(12): 1204005.
伍俊龙, 郭正华, 陈先锋, 等. 基于深度学习的光场成像三维测量方法研究 [J]. *中国激光*, 2020, 47(12): 1204005.
- [50] Guo M T, Zhu H, Zhou G Q, et al. Dense light field reconstruction from sparse sampling using residual network [M]//Jawahar C V, Li H D, Mori G, et al. *Computer vision-ACCV 2018. Lecture notes in computer science*. Cham: Springer, 2019, 11366: 50-65.
- [51] Vianello A, Ackermann J, Diebold M, et al. Robust hough transform based 3D reconstruction from circular light fields [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7327-7335.
- [52] Johannsen O, Sulc A, Goldluecke B. What sparse light field coding reveals about scene structure [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 3262-3270.
- [53] Wanner S, Goldluecke B. Globally consistent depth labeling of 4D light fields [C]//2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI, USA. New York: IEEE Press, 2012: 41-48.
- [54] 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), August 20-24, 2018, Beijing, China. New York: IEEE Press, 2018: 2362-2367.
- [55] 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, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 4748-4757.
- [56] Zhu H. Spatio-angular consistency for 4D light field processing [D]. Xi'an: Northwestern Polytechnical University, 2020.
朱昊. 基于空间角度一致性的四维光场处理方法研究 [D]. 西安: 西北工业大学, 2020.
- [57] 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.
- [58] Zhu H, Guo M T, Li H D, et al. Revisiting spatio-angular trade-off in light field cameras and extended applications in super-resolution [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2021, 27(6): 3019-3033.

- [59] 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.
- [60] Xie S L, Wang P, Sang X Z, et al. Augmented reality three-dimensional display with light field fusion [J]. *Optics Express*, 2016, 24(11): 11483-11494.
- [61] Kim H S, Jeong K M, Hong S I, et al. Analysis of image distortion based on light ray field by multi-view and horizontal parallax only integral imaging display [J]. *Optics Express*, 2012, 20(21): 23755-23768.
- [62] Isaksen A, McMillan L, Gortler S J. Dynamically reparameterized light fields [C] *Proceedings of the 27th annual conference on Computer graphics and interactive techniques-SIGGRAPH'00*, July 1, 2000, New York, USA. New York: ACM Press, 2000: 297-306.
- [63] Ng R, Levoy M, Brédif M, et al. Light field photography with a hand-held plenoptic camera [EB/OL]. [2021-05-30]. https://www.researchgate.net/publication/246688556_Light_Field_Photography_with_a_Hand-Held_Plenopic_Camera.
- [64] Levin A, Durand F. Linear view synthesis using a dimensionality gap light field prior [C] // *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE Press, 2010: 1831-1838.
- [65] Lin H T, Chen C, Kang S B, et al. Depth recovery from light field using focal stack symmetry [C] // *2015 IEEE International Conference on Computer Vision (ICCV)*, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 3451-3459.
- [66] Huang Z Y, Fessler J A, Norris T B, et al. Light-field reconstruction and depth estimation from focal stack images using convolutional neural networks [C] // *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 4-8, 2020, Barcelona, Spain. New York: IEEE Press, 2020: 8648-8652.
- [67] Liu C, Qiu J, Jiang M. Light field reconstruction from projection modeling of focal stack [J]. *Optics Express*, 2017, 25(10): 11377-11388.
- [68] Herman G T. *Fundamentals of computerized tomography: image reconstruction from projections* [M]. London: Springer, 2009.
- [69] Pan X C, Sidky E Y, Vannier M. Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction? [J]. *Inverse Problems*, 2009, 25(12): 1230009.
- [70] Landweber L. An iteration formula for Fredholm integral equations of the first kind [J]. *American Journal of Mathematics*, 1951, 73(3): 615-624.
- [71] Valencia S A, Rodriguez-Dagnino R M. Synthesizing stereo 3D views from focus cues in monoscopic 2D images [J]. *Proceedings of SPIE*, 2003, 5006: 377-388.
- [72] Chen Y C, Wu Y C, Liu C H, et al. Depth map generation based on depth from focus [C] // *2010 International Conference on Electronic Devices, Systems and Applications*, April 11-14, 2010, Kuala Lumpur, Malaysia. New York: IEEE Press, 2010: 59-63.
- [73] Malik A S, Shim S O, Choi T S. Depth map estimation using a robust focus measure [C] // *2007 IEEE International Conference on Image Processing*, September 16-October 19, 2007, San Antonio, TX, USA. New York: IEEE Press, 2007: 564-567.
- [74] Nayar S K. Shape from focus system for rough surfaces [J]. *Physics-Based Vision: Principles and Practice: Radiometry*, 1993, 1: 347-360.
- [75] Nayar S K, Nakagawa Y. Shape from focus [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994, 16(8): 824-831.
- [76] Strecke M, Alperovich A, Goldluecke B. Accurate depth and normal maps from occlusion-aware focal stack symmetry [C] // *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2529-2537.
- [77] Rajagopalan A N, Chaudhuri S, Mudenagudi U. Depth estimation and image restoration using defocused stereo pairs [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(11): 1521-1525.
- [78] Takahashi K, Kobayashi Y, Fujii T. From focal stack to tensor light-field display [J]. *IEEE Transactions on Image Processing*, 2018, 27(9): 4571-4584.
- [79] He K, Wang X L, Wang Z W, et al. Snapshot multifocal light field microscopy [J]. *Optics Express*, 2020, 28(8): 12108-12120.
- [80] Liu J, Xu T, Yue W, et al. Light-field moment microscopy with noise reduction [J]. *Optics Express*, 2015, 23(22): 29154-29162.
- [81] Liu F L, Kuo G, Antipa N, et al. Fourier DiffuserScope: single-shot 3D Fourier light field microscopy with a diffuser [J]. *Optics Express*, 2020, 28(20): 28969-28986.
- [82] Kim J, Jung J H, Jeong Y, et al. Real-time integral imaging system for light field microscopy [J]. *Optics*

- Express, 2014, 22(9): 10210-10220.
- [83] Ren X F, Malik J. Learning a classification model for segmentation[C]//Proceedings Ninth IEEE International Conference on Computer Vision, October 13-16, 2003, Nice, France. New York: IEEE Press, 2003: 10-17.
- [84] Zhu H, Zhang Q, Wang Q. 4D light field superpixel and segmentation [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6709-6717.
- [85] Khan N, Zhang Q, Kasser L, et al. View-consistent 4D light field superpixel segmentation [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea. New York: IEEE Press, 2019: 7810-7818.
- [86] Zhu H, Sun X M, Zhang Q, et al. Full view optical flow estimation leveraged from light field superpixel [J]. IEEE Transactions on Computational Imaging, 2020, 6: 12-23.
- [87] Chen J, Hou J H, Ni Y, et al. Accurate light field depth estimation with superpixel regularization over partially occluded regions[J]. IEEE Transactions on Image Processing, 2018, 27(10): 4889-4900.
- [88] Su X, Maugey T, Guillemot C. Rate-distortion optimized graph-based representation for multiview images with complex camera configurations [J]. IEEE Transactions on Image Processing, 2017, 26(6): 2644-2655.
- [89] Karypis G, Kumar V. Multilevel k -way hypergraph partitioning[J]. VLSI Design, 2000, 11(3): 285-300.
- [90] Rizkallah M, Su X, Maugey T, et al. Geometry-aware graph transforms for light field compact representation [J]. IEEE Transactions on Image Processing, 2020, 29: 602-616.
- [91] Lü X, Wang X, Wang Q, et al. 4D light field segmentation from light field super-pixel hypergraph representation[J]. IEEE Transactions on Visualization and Computer Graphics, 2020, 2158: 32203022.
- [92] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(11): 1222-1239.
- [93] Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(9): 1124-1137.
- [94] Maugey T, Ortega A, Frossard P. Graph-based representation for multiview image geometry [J]. IEEE Transactions on Image Processing, 2015, 24(5): 1573-1586.
- [95] Su X, Rizkallah M, Maugey T, et al. Graph-based light fields representation and coding using geometry information[C]//2017 IEEE International Conference on Image Processing (ICIP), September 17-20, 2017, Beijing, China. New York: IEEE Press, 2017: 4023-4027.
- [96] Rizkallah M, Su X, Maugey T, et al. Graph-based transforms for predictive light field compression based on super-pixels [C] // 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), April 15-20, 2018, Calgary, AB, Canada. New York: IEEE Press, 2018: 1718-1722.
- [97] Zhou T H, Tucker R, Flynn J, et al. Stereo magnification: learning view synthesis using multiplane images[J]. ACM Transactions on Graphics, 2018, 37(4): 65.
- [98] Srinivasan P P, Tucker R, Barron J T, et al. Pushing the boundaries of view extrapolation with multiplane images [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 175-184.
- [99] Mildenhall B, Srinivasan P P, Ortiz-Cayon R, et al. Local light field fusion [J]. ACM Transactions on Graphics, 2019, 38(4): 29.
- [100] Flynn J, Broxton M, Debevec P, et al. DeepView: view synthesis with learned gradient descent [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 2362-2371.
- [101] Li Q B, Kalantari N K. Synthesizing light field from a single image with variable MPI and two network fusion[J]. ACM Transactions on Graphics, 2020, 39(6): 229.
- [102] Tucker R, Snavely N. Single-view view synthesis with multiplane images [C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 548-557.
- [103] Bracewell R N. Strip integration in radio astronomy [J]. Australian Journal of Physics, 1956, 9(2): 198.
- [104] Levoy M. Volume rendering using the Fourier projection-slice theorem [M]. San Francisco: Morgan

- Kaufmann Publishers Inc, 1992.
- [105] Ng R. Fourier slice photography[J]. ACM Transactions on Graphics, 2005, 24(3): 735-744.
- [106] Macovski A. Medical imaging systems[M]. London: Prentice Hall, 1983.
- [107] Deans S R. The Radon transform and some of its applications[M]. New York: John Wiley and Sons Inc, 1983.
- [108] Malzbender T. Fourier volume rendering[J]. ACM Transactions on Graphics, 1993, 12(3): 233-250.
- [109] Zhao Q Q, Zhang T, Zheng W B. Research on high resolution digital refocusing of light field imaging based on microlens array[J]. Laser & Optoelectronics Progress, 2016, 53(10): 101001.
赵青青, 张涛, 郑伟波. 基于微透镜型光场成像的高分辨率数字对焦技术研究[J]. 激光与光电子学进展, 2016, 53(10): 101001.
- [110] Levin A, Hasinoff S W, Green P, et al. 4D frequency analysis of computational cameras for depth of field extension[J]. ACM Transactions on Graphics, 2009, 28(3): 1-14.
- [111] 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): 1-13.
- [112] Levoy M, Ng R, Adams A, et al. Light field microscopy [J]. ACM Transactions on Graphics, 2006, 25(3): 924-934.
- [113] Le Pendu M, Guillemot C, Smolic A. A Fourier disparity layer representation for light fields [J]. IEEE Transactions on Image Processing, 2019, 28(11): 5740-5753.
- [114] Le Pendu M, Smolic A. High resolution light field recovery with Fourier disparity layer completion, demosaicing, and super-resolution [C] // 2020 IEEE International Conference on Computational Photography (ICCP), April 24-26, 2020, St. Louis, MO, USA. New York: IEEE Press, 2020: 19653339.
- [115] Dib E, Le Pendu, Guillemot C. Light field compression using Fourier disparity layers [C] // 2019 IEEE International Conference on Image Processing (ICIP), September 22-25, 2019, Taipei, Taiwan, China. New York: IEEE Press, 2019: 3751-3755.
- [116] Le Pendu M, Ozcinar C, Smolic A. Hierarchical Fourier disparity layer transmission for light field streaming [C] // 2020 IEEE International Conference on Image Processing (ICIP), October 25-28, 2020, Abu Dhabi, United Arab Emirates. New York: IEEE Press, 2020: 2606-2610.
- [117] Li Y N, Wang X, Zhou G Q, et al. Deep anti-aliasing of whole focal stack using its slice spectrum [EB/OL]. (2021-01-23) [2021-05-30]. <https://arxiv.org/abs/2101.09420>.