http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0797

面向低空场景的光场高斯泼溅表示及空间-角度同时超分辨率方法

孟佳涵[1]，杨驰[1]，李传君[2]，林丽莉[2]，周文晖[1,*]
(1. 杭州电子科技大学 计算机学院，杭州 310018；2. 浙江工商大学 信息与电子工程学院，杭州 310018)

摘 要：光场成像技术可同时捕获光线的空间位置和角度信息，在无人机感知、环境监测等低空经济应用场景中具有广阔应用前景。然而，现有光场相机成像及超分辨率方法难以兼顾空间分辨率与角度分辨率，制约了其在精确感知和高精度建模等任务中的应用。为此，提出一种面向低空场景的光场高斯泼溅表示及光场空间-角度同时超分辨率（LFASSR）方法。将传统3D高斯泼溅建模方法拓展到4D光场表示空间中，提出一种4D光场高斯泼溅表示方法，并利用光场极平面（EPI）的空间-角度几何特性，将其分解为水平与垂直方向上的2D高斯泼溅级联建模，以更有效地表征光场几何特征。基于此，构建了一个通用的光场图像超分辨率框架，在提取并融合光场空间、角度及EPI域特征后，构建水平和垂直方向的光场几何特征空间，并分别进行2D高斯泼溅建模，最终经EPI渲染，生成致密视角且高空间分辨率的高质量光场图像。在公开光场图像数据集及低空场景光场数据上的实验结果表明：所提方法在LFASSR任务中显著优于现有方法，在峰值信噪比（PSNR）和结构相似性指标（SSIM）等客观指标上均达到当前最优水平，在视觉质量上展现出清晰的细节恢复能力与自然的视差过渡效果。在独立的光场空间超分辨率（LFSSR）任务和光场角度超分辨率（LFASR）任务中，所提方法同样展现出良好的泛化能力与竞争力。

关 键 词：光场；光场超分辨率；光场高斯泼溅；空间-角度同时超分辨率；极平面几何
中图分类号：TP391.4
文献标志码：A **文章编号**：1001-5965(2026)09-3189-13

随着低空经济的快速发展，无人机巡检、城市空中交通、精准农业和三维环境感知等新兴应用场景对多维视觉信息的需求日益增长。相较于传统的二维成像机制，光场成像技术不仅记录了场景中光线的强度，还保留了其传播方向信息，因而能在单次拍摄中捕获多视角视差、透视偏移、深度线索在内的高维视觉信息[1-2]。这一特性使其在实现重聚焦、视点合成与三维感知等任务中具有独特优势，正逐步发展为支撑低空智能感知系统的关键技术之一。

然而，受限于微透镜阵列的物理采样机制，现有光场成像在空间分辨率与角度分辨率之间存在固有的权衡矛盾[3]，导致成像质量难以满足低空经济应用中高精度感知与精细化建模的任务需求，制约了光场技术在相关场景中的应用。为突破光场相机的物理成像极限，近年来光场超分辨率技术受到广泛关注，并取得显著进展。

光场超分辨率技术主要分为光场角度超分辨率（light field angular super-resolution, LFASR）和光场空间超分辨率（light field spatial super-resolution,

收稿日期：2025-11-17；录用日期：2025-12-12；网络出版时间：2026-02-02 11:35
网络出版地址：link.cnki.net/urlid/11.2625.V.20260130.1851.004
基金项目：国家自然科学基金 (62471437)；浙江省“尖兵”“领雁”科技计划项目 (2023C01030)
*通信作者. E-mail：zhouwenhui@hdu.edu.cn

引用格式：孟佳涵，杨驰，李传君，等. 面向低空场景的光场高斯泼溅表示及空间-角度同时超分辨率方法 [J]. 北京航空航天大学学报，2026，52（9）：3189-3201. Meng J H，Yang C，Li C J，et al. Light field gaussian splatting representation and joint spatial-angular super-resolution for low-altitude scenarios [J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（9）：3189-3201（in Chinese）.

LFSSR)2类。其中,LFASR聚焦于光场的光线角度域,增加光线角度的采样密度,提升光场图像在角度域上的信息丰富度;LFSSR旨在提升光场子孔径图像的空间分辨率,以增强光场图像在空间维度上的细节表现。

在LFASR研究中,Kalantari等[4]首次将深度学习引入该任务,将新视角合成分解为视差估计和颜色估计2个阶段,并通过串联的卷积神经网络(convolutional neural network, CNN)分别建模。随后,Zhang等[5]利用垂直和水平光场极平面(epipolar plane image, EPI)特征,通过反卷积实现角度域致密采样。针对大基线稀疏光场,Jin等[6]基于光场几何设计了深度估计与新视角合成的融合框架。Liu等[7]在宏像素特征空间进行LFASR,在有效处理大视差的同时降低推理开销。Wang等[8]提出的光场特征卷积在保留视角相关性与EPI结构的同时缓解空间与角度信息不平衡问题。Mao等[9]通过自适应对齐和特征解耦,提高了模型对复杂场景和低光照的适应能力,减少了新视角合成中的视觉伪影。

在LFSSR研究中,Yuan等[10]先对各子孔径图像独立进行LFSRR,再在EPI域通过卷积模块增强几何一致性。Wang等[11]提出双向循环卷积网络,分别建模水平和垂直方向相邻子孔径图像间的空间关系,并通过堆叠泛化进行融合;在此基础上,Zhang等[12]引入对角线方向建模,并采用残差融合模块实现光场的高分辨空间重建。随后,Liang等[13]通过角度Transformer和空间Transformer建模视角间互补信息及子孔径图像内的长短程依赖。Gao等[14]与Chen等[15]分别基于状态空间模型引入Mamba架构和选择性扫描机制,在EPI或空间-角度联合特征建模中实现了更优的性能与效率平衡。

尽管光场超分辨率技术已取得显著进展,但大多数现有研究都聚焦于LFSSR或LFASR的单一任务,不同任务所采用的网络结构亦存在差异,且难以实现光场图像的空间-角度同时超分辨率。为此,Wang等[16]设计了一种通用光场图像超分辨率框架Distg(SSR/ASR),以适用于LFSSR和LFASR任务。Van Duong等[17]探索了空间-角度域内的像素相关性,以重建空间或角度高分辨率的光场图像。Yoon等[18]尝试了光场图像的空间与角度同时超分辨率,分别训练了基于CNN的LFSRR网络和LFASR网络,然后,通过端到端方式进行联合微调。Liu等[19]面向光场空间-角度同时超分辨率(light field angular spatial super-resolution, LFASSR)的光场重建方法,通过自适应聚合多视角像素信息,在空间域与角度域中同时建模视角间的相关性,实现了致密高分辨率光场图像的联合合成。Liu等[20]进一步提出了一种隐式细节增强网络,在隐式表征空间中学习不同方向上的像素差分关系,有效提升了LFASSR任务中细节恢复的质量,重建出纹理更加丰富、结构更加一致的高分辨率光场图像。然而,与单一的LFSSR或LFASR任务相比,LFASSR研究相对较少,实现2个维度的同步增强仍面临挑战。

本文提出了一种面向低空场景的光场高斯泼溅表示及LFASSR框架。不同于传统光场超分方法通常在空间域或角度域分别建模、逐步重建的策略,本文将近年来新兴的3D高斯泼溅技术[21-22]引入光场建模中,并进一步扩展至光场的双平面表示空间,提出了4D光场高斯泼溅表示。该表示充分利用光场水平和垂直EPI的空间-角度几何特性,将其分解为水平与垂直方向上2个2D高斯泼溅的级联表示,避免了传统方法中空间与角度超分相互独立或顺序执行所带来的误差累积,从而在光场EPI域上实现真正意义的空间-角度同时高分辨率渲染。本文主要贡献如下:

1)设计一种基于4D光场高斯泼溅表示的光场图像超分辨率通用框架,不仅可实现光场图像空间域或角度域的单域超分辨率,而且可实现LFASSR。

2)提出一种4D光场高斯泼溅建模方法,利用光场EPI的空间-角度几何特性,将4D光场高斯泼溅表示分解为水平与垂直方向上2个2D高斯泼溅的级联表示,并将水平/垂直EPI中的像素点扩散为连续高斯场,通过渲染相互堆叠的高斯场,恢复高分辨率光场的几何细节。

3)提出的光场图像超分辨率方法已在多种光场图像超分辨率任务中进行了测试,包括LFASSR、单一的LFSSR和LFASR任务,在评价指标和视觉效果方面均优于现有方法;在低空场景光场数据上的泛化测试也展现出令人满意的性能和渲染效果。

1 本文方法

1.1 4D光场高斯泼溅表示

光场双平面表示法$\boldsymbol{L}(u,v,x,y)\in\mathbf{R}^{U\times V\times W\times H}$采用一个四维函数来描述场景中$(u,v)$视角方向上的光线在子孔径图像坐标$(x,y)$上的投影,$(u,v)$与$(x,y)$分别为场景中光线与光场角度域平面和空间域平面的交点坐标。

本文首先将传统的3D高斯泼溅技术扩展到光

场双平面表示中, 提出一种 4D 光场高斯泼溅(4D light field Gaussian splatting, 4DLFGS)表示方法。该方法将场景用一组 4D 高斯场表示, 每个高斯场的属性除均值向量外都可学习, 包括协方差矩阵、不透明度和颜色信息。具体而言, 令场景中第 i 条光线的双平面坐标为 (u_i, v_i, x_i, y_i), 假定其服从于高斯分布, 则该光线的 4D 高斯场定义为

$$f_i(\boldsymbol{p}) = \sigma_i \exp\left(-\frac{1}{2}(\boldsymbol{p}-\boldsymbol{\mu}_i)^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}(\boldsymbol{p}-\boldsymbol{\mu}_i)\right) \tag{1}$$

式中: $\boldsymbol{p} = (u, v, x, y)^{\mathrm{T}}$ 为光场双平面表示中的任意点位置; $\boldsymbol{\mu}_i = (u_i, v_i, x_i, y_i)^{\mathrm{T}}$ 为 4D 高斯场的均值向量; $\boldsymbol{\Sigma}_i \in \mathbf{R}^{4\times4}$ 为其协方差矩阵; σ_i 为不透明度; $f_i(\boldsymbol{p})$ 为第 i 个高斯场对 $\boldsymbol{p}$ 点的渲染权重。

在高斯泼溅渲染过程中, 采用传统 3D 高斯泼溅中的 α 混合技术, 双平面坐标 $\boldsymbol{p}$ 点的颜色信息 $C_{\mathrm{LFGS}}(\boldsymbol{p})$ 由所有 4D 高斯场加权得到:

$$C_{\mathrm{LFGS}}(\boldsymbol{p}) = \sum_{i\in N} c_i f_i^{\mathrm{2D}}(\boldsymbol{p}) \prod_{j=1}^{i-1}\left(1 - f_j^{\mathrm{2D}}(\boldsymbol{p})\right) \tag{2}$$

式中: $f_i^{\mathrm{2D}}(\cdot)$ 为 $f_i(\cdot)$ 在二维平面的投影; c_i 为第 i 个高斯场的 RGB 值或球谐函数系数。

根据光场水平和垂直 EPI 表示及其空间-角度几何特性, 本文进一步将 4D 光场高斯泼溅建模分解为水平和垂直方向上 2 个 2D 高斯泼溅的级联建模。具体而言, 将光场双平面坐标 (u, v, x, y) 划分为水平 EPI 坐标 (u, x) 和垂直 EPI 坐标 (v, y), 从而将 4D 高斯模型 $f_i(\boldsymbol{p})$ 分解为水平 EPI 上的 2D 高斯模型 $f_{i,u}(\boldsymbol{p}_u)$ 和垂直 EPI 上的 2D 高斯模型 $f_{i,v}(\boldsymbol{p}_v)$ 的级联形式, 如式 (3) 所示。

$$\begin{cases} f_i(\boldsymbol{p}) = f_{i,u}(\boldsymbol{p}_u)\cdot f_{i,v}(\boldsymbol{p}_v) \\ f_{i,u}(\boldsymbol{p}_u) = \sigma_{i,u}\exp\left(-\frac{1}{2}(\boldsymbol{p}_u-\boldsymbol{\mu}_{i,u})^{\mathrm{T}}\boldsymbol{\Sigma}_{i,u}^{-1}(\boldsymbol{p}_u-\boldsymbol{\mu}_{i,u})\right) \\ f_{i,v}(\boldsymbol{p}_v) = \sigma_{i,v}\exp\left(-\frac{1}{2}(\boldsymbol{p}_v-\boldsymbol{\mu}_{i,v})^{\mathrm{T}}\boldsymbol{\Sigma}_{i,v}^{-1}(\boldsymbol{p}_v-\boldsymbol{\mu}_{i,v})\right) \end{cases} \tag{3}$$

式中: $\boldsymbol{p}_u = (u, x)^{\mathrm{T}}$ 和 $\boldsymbol{p}_v = (v, y)^{\mathrm{T}}$ 分别为水平和垂直 EPI 坐标; $\boldsymbol{\mu}_{i,u} = (u_i, x_i)^{\mathrm{T}}$ 和 $\boldsymbol{\mu}_{i,v} = (v_i, y_i)^{\mathrm{T}}$ 分别为水平 EPI 和垂直 EPI 上 2D 高斯模型的均值向量; $\boldsymbol{\Sigma}_{i,u}$ 和 $\boldsymbol{\Sigma}_{i,v}$ 为对应的协方差矩阵。

进而, 4D 光场高斯泼溅的渲染过程分解为 2 个 2D 高斯泼溅的级联渲染, 如式 (4) 所示。

$$\begin{cases} C_{\mathrm{LFGS}}(\boldsymbol{p}_v) = \sum_{j\in N} c_{j,v} f_{j,v}^{\mathrm{2D}}(\boldsymbol{p}_v) \prod_{j=1}^{j-1}\left(1 - f_{j,v}^{\mathrm{2D}}(\boldsymbol{p}_v)\right) \\ c_v(\boldsymbol{p}_u) = \sum_{i\in N} c_{i,u} f_{i,u}^{\mathrm{2D}}(\boldsymbol{p}_u) \prod_{i=1}^{i-1}\left(1 - f_{i,u}^{\mathrm{2D}}(\boldsymbol{p}_u)\right) \end{cases} \tag{4}$$

式中: $f_{i,u}^{\mathrm{2D}}(\cdot)$ 和 $f_{j,v}^{\mathrm{2D}}(\cdot)$ 分别为 $f_{i,u}(\cdot)$ 和 $f_{j,v}(\cdot)$ 的 2D 平面投影; $c_{i,u}$ 和 $c_{j,v}$ 分别为水平和垂直 EPI 中的高斯模型颜色信息。为减少计算量, 具体实现时采用简化 α 混合技术, 即式 (4) 简化为

$$\begin{cases} C_{\mathrm{LFGS}}(\boldsymbol{p}_v) = \sum_{j\in N} c_{j,v} f_{j,v}^{\mathrm{2D}}(\boldsymbol{p}_v) \\ c_v(\boldsymbol{p}_u) = \sum_{i\in N} c_{i,u} f_{i,u}^{\mathrm{2D}}(\boldsymbol{p}_u) \end{cases} \tag{5}$$

基于上述分解过程, 本文采用水平和垂直 EPI 上的 2 个级联 2D 高斯泼溅模型来表示光场在水平与垂直 EPI 上的几何结构, 从而实现对四维光场整体几何分布的高效表示。

1.2　整体网络结构

本文提出的 LFASSR 网络架构包含 3 个模块: 光场几何编码模块、4D 光场高斯泼溅模块和光场图像生成模块, 实现了对输入光场图像 $\boldsymbol{L}(u, v, x, y) \in \mathbf{R}^{U\times V\times W\times H}$ 在其空间和角度域上的同时超分辨率, 得到空间-角度高分辨率的光场图像 $\tilde{\boldsymbol{L}}(u, v, x, y) \in \mathbf{R}^{U'\times V'\times W'\times H'}$。

如图 1 所示, 首先, 以一个稀疏视角且低空间分辨率的光场图像 $\boldsymbol{L}(u, v, x, y)$ 作为输入, 在光场几何编码模块中提取和融合光场图像的空间、角度和 EPI 特征, 得到空间-角度联合特征 $\boldsymbol{Z}_{\mathrm{g}}$; 然后, 将 $\boldsymbol{Z}_{\mathrm{g}}$ 输入 4D 光场高斯泼溅模块中建模高斯场, 并在光场空间和角度域上实现上采样渲染, 得到光场潜在特征编码 $\boldsymbol{Z}'$; 最后, 由光场图像生成模块从潜在特征编码 $\boldsymbol{Z}'$ 中生成空间-角度高分辨率的光场图像 $\tilde{\boldsymbol{L}}(u, v, x, y)$。

本文在模型训练阶段使用 $L1$ 损失函数, 以衡量生成的高分辨率光场图像 $\tilde{\boldsymbol{L}}$ 与真实高分辨率光场图像 $\hat{\boldsymbol{L}}$ 间的像素级差异, 通过最小化损失函数来优化模型参数, 具体损失函数为

$$\mathcal{L} = \sum \left|\tilde{\boldsymbol{L}}(u, v, x, y) - \hat{\boldsymbol{L}}(u, v, x, y)\right| \tag{6}$$

1.3　光场几何编码模块

光场几何编码模块旨在从输入光场图像中提取其空间、角度及 EPI 的联合特征充分挖掘光场图像中的空间/几何一致性, 以及视角关联性, 为后续超分辨率重建提供高阶特征的表征能力。具体如图 2(a) 所示, 输入的光场图像 $\boldsymbol{L}(u, v, x, y)$ 先经过一组由 3×3 卷积层组成的卷积模块, 逐层提升特征通道数至 64, 从而得到初始特征图 $\boldsymbol{F}_{\mathrm{init}}$。随后, 引入一个自适应特征融合模块[23], 该模块以分块形式逐级融合光场几何特征, 在每块中使用不同卷积层分别提取空间、角度、垂直 EPI 和水平 EPI 的特征, 并在每块的末端采用短残差连接, 以稳定梯度传播并有效保留结构相关的浅层特征。

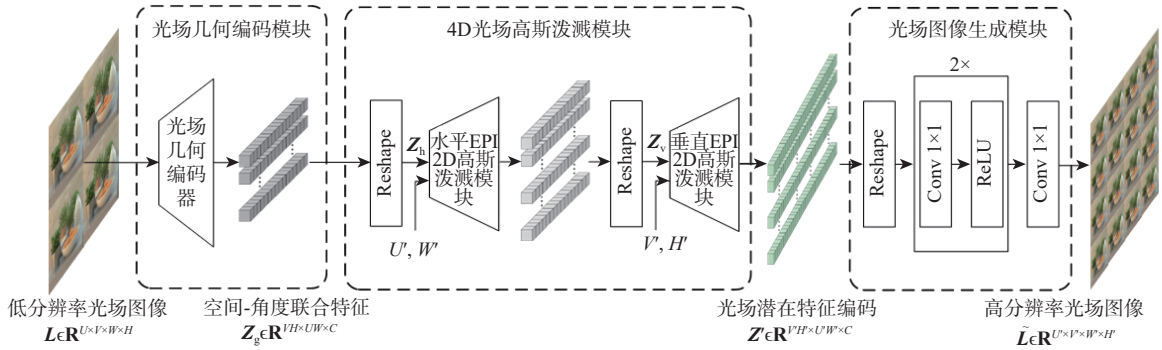
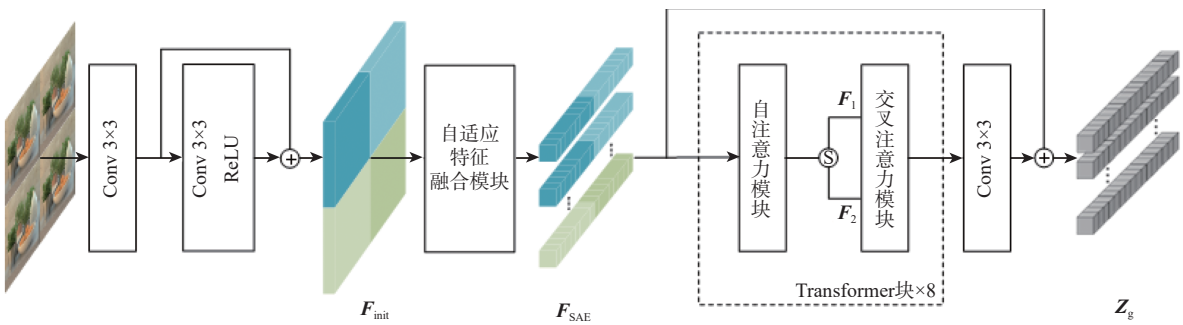
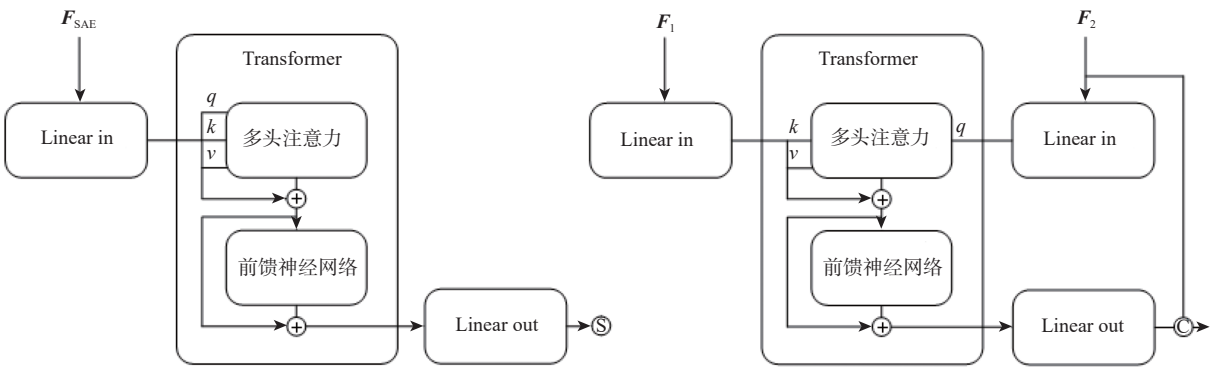


图 1 整体网络结构框

Fig. 1 Overall framework of the proposed network



(a) 光场几何编码模块



(b) 自注意力模块 (c) 交叉注意力模块

图 2 光场几何编码模块和注意力模块

Fig. 2 Light field geometry encoding module and attention modules

将经过残差网络训练增强得到的特征 $\boldsymbol{F}_{\mathrm{SAE}}$ 输入一个基于 Transformer 的特征增强单元。该单元利用注意力机制捕获长程依赖关系，从全局视角整合光场不同视角间的结构信息，以获得更完整且表达力更强的空间-角度联合特征 $\boldsymbol{Z}_{\mathrm{g}}$。该特征增强单元由 8 个 Transformer 块组成，每块都包含自注意力模块和交叉注意力模块。每个自注意力/交叉注意力模块内部都包含一个多头注意力结构和一个前馈神经网络，如图 2(b) 和图 2(c) 所示。线性输入层（Linear in）和线性输出层（Linear out）分别用来提升和压缩特征通道数，在保证输入/输出维度一致的同时增强特征表示能力。在自注意力模块中，直接将整个光场特征图作为多头注意力机制的输入，以捕获特征间的全局依赖关系并增强特征表达能力。

在交叉注意力模块中，输入特征首先在通道维度上进行分割操作，得到 2 个特征分支 $\boldsymbol{F}_1$ 和 $\boldsymbol{F}_2$；随后将 $\boldsymbol{F}_1$ 和 $\boldsymbol{F}_2$ 输入到 2 个独立训练的线性输入层中以提升通道维度，并将其分别作为键值对 (k, v) 和查询 (q) 输入多头注意力结构中，并经前馈神经网络增强非线性表示能力。将线性输出层生成的特征与输入特征 $\boldsymbol{F}_2$ 进行张量拼接，传递至下一个 Transformer 块。

经所有 Transformer 块的特征提取与全局建模后，输出特征被送入一个 3×3 卷积层，以补充 Transformer 结构在细节层面上建模能力不足的问题，从而增强特征的局部一致性与纹理保真度。随后，将该卷积层的输出与 Transformer 特征增强单元的输入 $\boldsymbol{F}_{\mathrm{SAE}}$ 通过短残差连接进行相加，在融合局部

与全局信息的同时保持特征分布的稳定性, 从而获得几何一致性的光场特征表示。

1.4　4D 光场高斯泼溅模块

4D 光场高斯泼溅模块旨在在光场特征空间中构建 1.1 节所述的 4D 光场高斯泼溅模型, 并将其分解为垂直和水平 EPI 特征图上的 2 个 2D 高斯泼溅模块, 再在高分辨率网格上对 EPI 特征图进行上采样渲染。具体如图 1 所示, 空间-角度联合特征 $\boldsymbol{Z}_{\mathrm{g}}$ 先经过 Reshape 操作得到水平 EPI 特征图 $\boldsymbol{Z}_{\mathrm{h}}$, 维度为 (VHC,U,W); 将其输入到水平 EPI 2D 高斯泼溅模块中, 对角度维度 U 和空间维度 W 进行上采样和渲染, 得到增强后的水平 EPI 特征图 $\boldsymbol{Z}'_{\mathrm{h}}$, 维度为 (VHC,U',W')。随后, $\boldsymbol{Z}'_{\mathrm{h}}$ 经过 Reshape 操作将超分后的 U'、W' 和特征通道 C 合并为一个维度, 得到垂直 EPI 特征图 $\boldsymbol{Z}_{\mathrm{v}}$, 维度为 $(U'W'C,V,H)$, 并送入垂直 EPI 2D 高斯泼溅模块完成对 V 和 H 维度的上采样和渲染, 得到增强后维度为 $(U'W'C,V',H')$ 的垂直 EPI 特征图 $\boldsymbol{Z}'_{\mathrm{v}}$, 重组为高维光场特征编码 $\boldsymbol{Z}'\in\mathbf{R}^{V'H'\times U'W'\times C}$。

EPI 特征图的 2D 高斯泼溅渲染示意如图 3 所示, 低分辨率特征网格上的每个特征点(图中红实心点)都被建模为一个 2D 高斯椭圆(场), 则高分辨率网格中特征点 $\boldsymbol{u}_{\mathrm{q}}$ 的颜色信息受到周围所有 2D 高斯椭圆的叠加影响。根据式 (3), 第 i 个 2D 高斯椭圆对 $\boldsymbol{u}_{\mathrm{q}}$ 渲染的权重值可表示为

$$f_i(\boldsymbol{u}_{\mathrm{q}})=\sigma_i\exp\left(-\frac{1}{2}(\boldsymbol{u}_{\mathrm{q}}-\boldsymbol{u}_i)^{\mathrm{T}}\boldsymbol{\Sigma}_i^{-1}(\boldsymbol{u}_{\mathrm{q}}-\boldsymbol{u}_i)\right)\tag{7}$$

式中: 均值向量 $\boldsymbol{u}_i\in\mathbf{R}^2$ 为低分辨率 EPI 特征图中第 i 个特征点坐标; 协方差矩阵 $\boldsymbol{\Sigma}_i$ 和 σ_i 分别为第 i 个 2D 高斯椭圆的形状参数和不透明度。

训练时, 先在低分辨率网格上初始化每个 2D 高斯椭圆, 并在训练过程中不断优化光场 EPI 特征图、协方差矩阵和不透明度。将所有 2D 高斯椭圆作用在高分辨率网格上, 生成高分辨率 EPI 特征图。根据式 (5), 高分辨率网格中特征点 $\boldsymbol{u}_{\mathrm{q}}$ 的颜色信息可由式 (8) 计算得到:

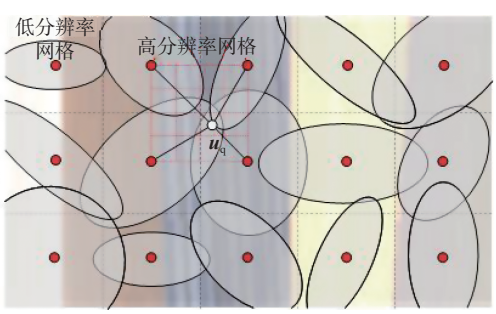


图 3　低分辨率 EPI 网格上构建的 2D 高斯泼溅模型示意

Fig. 3　Illustration of the 2D Gaussian splatting model constructed on a low-resolution EPI grid

$$\boldsymbol{Z}_{\mathrm{r}}(\boldsymbol{u}_{\mathrm{q}})=\sum_{i\in N}f_i(\boldsymbol{u}_{\mathrm{q}})\boldsymbol{z}_i\tag{8}$$

式中: $\boldsymbol{z}_i$ 为低分辨率 EPI 特征图中第 i 个特征点的特征向量。

1.5　光场图像生成模块

在完成 4D 光场高斯泼溅建模和渲染后, 还需将高维光场特征编码 $\boldsymbol{Z}'$ 还原为相应的光场双平面表示, 并优化生成结果。为此, 本文设计了一个光场图像生成模块, 用于将高斯泼溅渲染后的特征映射回光场双平面表示。具体而言, 该模块将 $\boldsymbol{Z}'$ 作为输入, 首先, 采用 Reshape 操作将光场特征图的维度转化为 $(C,U'V',H',W)$。然后, 依次通过 3 个连续的 1×1 卷积层, 将特征图的通道数 C 逐步压缩至单通道, 以恢复图像的光强信息。每个卷积层之间均引入 LeakyReLU 激活函数, 以增强网络的非线性建模能力。最后, 将卷积层输出结果重组为光场子孔径图像阵列, 从而得到最终的高分辨率光场图像 $\tilde{\boldsymbol{L}}$。

2　实　验

对比实验首先采用当前 LFASR/LFSSR 任务的 2 套常用公开数据集。其中, LFASR 数据集包括 2 个合成场景光场数据集 HCInew 和 HCIold 及真实场景光场数据集 30Scenes、Occlusion 和 Reflective。LFSSR 数据集包括合成场景数据集 HCInew 和 HCIold 及真实场景数据集 EPFL、INRIA 和 STFgantry。训练集与测试集的划分方式与 Distg[16] 保持一致。在面向低空场景的泛化性能测试中, 从 UrbanLF-Real[24] 光场数据集共 164 个测试场景中选取 12 个典型低空应用的光场图像作为测试集, 这些光场图像由第二代光场相机 Lytro Illum 采集得到, 包含车辆、道路、建筑等低空场景。图 4 展示了其中部分场景示例。在上述数据集中完成了下列实验: ①在 LFASSR 任务中, 将角度分辨率从 2 像素×2 像素放大至 5×5, 同时放大空间分辨率为原来的 2 倍; ②在单一的 LFSSR 任务中, 保持角度分辨率不变, 将 5 像素×5 像素光场图像的空间分辨率提升至 2 倍或 4 倍; ③在单一的 LFASR 任务中, 保持空间分辨率不变, 将角度分辨率从 2 像素×2 像素放大至 7 像素×7 像素; ④在低空场景测试集上测试 LFASSR 模型的泛化性能。比对方法选取近年来表现优异且具有代表性的模型作为参考。在实验中, 所有选取的对比方法均在相同的光场数据集上进行训练和测试。对于 LFASR 和 LFSSR 任务, 在表中直接以原始论文提供的结果作为对比数据。而在 LFASSR 中, 实验选取一些提供了源码的角度或空间超分模型进行组合, 并在同样的数据集下进行重新训练。



图4 低空场景(建筑、桥梁、交通、道路)下的光场图像及其EPI图像

Fig. 4 Light field images and their EPIs in low-altitude scenarios (e.g., buildings, bridges, traffic, and roads)

2.1 实验细节

本文实验在一台配备Intel Xeon E5-2680 v4 CPU、64 GB DDR4内存和一张NVIDIA RTX 3090 GPU的服务器上进行。网络训练采用自适应矩估计(Adam)优化器,其中,一阶矩估计的衰减率$\beta_1=0.9$,二阶矩估计的衰减率$\beta_2=0.999$。每次训练的批次大小为1。初始学习率设为1×10^{-4},学习率每15个训练轮次减半,模型在训练约70轮后基本收敛。

在指标评估时,由于现有光场超分辨率方法均在YCbCr色彩空间中的Y通道上进行评估,因此,本文也同样在光场图像的Y通道计算峰值信噪比(peak signal-to-noise ratio, PSNR)和结构相似性指标(structural similarity index, SSIM)等指标。具体而言,本文在所有子孔径图像上计算超分辨率重建结果与Ground Truth之间的PSNR和SSIM,并取其平均值作为光场重建质量的评价指标。值得注意的是,在LFASR任务中,输入的2×2光场在角度维度的超分过程中不涉及插值操作,因此,子孔径阵列4个角位置的图像与Ground Truth完全一致,其指标显著高于其余子孔径图像。为避免该现象对整体评估结果造成干扰,本文在计算平均PSNR和SSIM时排除了这4个角上的子孔径图像。相比之下,在LFSSR任务及空间-角度联合超分辨率任务中,所有子视图均参与了插值和重建过程,因此,在指标计算时纳入全部子孔径图像。另外,在生成高分辨率的RGB光场图像时,本文先将低分辨率RGB光场图像转换至YCbCr颜色空间,并将Y通道(亮度信息)作为模型的输入。对于Cb和Cr通道(色度信息),采用双三次插值进行上采样。将模型生成的高分辨率Y通道与上采样后的Cb、Cr通道重新合并,经过转换得到高分辨率的RGB光场图像。

处理光场数据集时,首先提取原始光场图像中心的5像素×5像素或7像素×7像素作为Ground Truth光场图像(由光场超分辨率任务确定),再从中提取位于子孔径图像阵列四个角上的子孔径图像,构建角度分辨率为2像素×2像素的稀疏光场,作为LFASR任务的输入。对于所有超分辨率任务,在生成训练集时,都需将原始光场图像分割成小像素块。在LFSSR和LFASSR任务中,还需将初始像素块下采样到32像素×32像素作为低分辨率输入。对于空间分辨率放大2倍和4倍的任务,初始像素块的空间分辨率分别是64像素×64像素和128像素×128像素。

为提升模型泛化能力,在训练过程中对光场图像进行水平、垂直及对角线方向的翻转操作。在复现其他方法时,同样采用相同的数据增强策略,以确保实验设置的一致性与公平性。

2.2 LFASSR任务

在LFASSR任务中,由于现有的研究缺乏LFASSR模型,本实验将现有单一空间和单一LFASR模型组合起来进行训练,一共对比了6种拼接模型和1种同时超分辨率模型LFSR[19]。所有拼接模型都是先LFASR再LFSSR。数据集采用LFASR任务中3个真实光场数据集(30Scenes、Occlusion和Reflective)及LFSSR任务中的EPFL、HCInew、HCIold数据集。表1评估了光场图像在空间分辨率上放大2倍,角度分辨率从2像素×2像素放大至5像素×5像素的合成质量。量化比对结果表明,本文模型的PSNR和SSIM指标均优于其他模型[25]。

图5比对了各模型在3个测试场景中的超分辨率视觉效果,本文模型的整体重建效果更接近真实光场图像。在ISO_Chart和Reflective_22场景中,本文模型在生成细密纹理方面表现更为出色。在Occlusions_35场景中,现有模型的生成结果在一些区域出现颜色缺失(黄框标记),EPI图像也可观察到这些现象。

2.3 LFASR任务

在单一LFASR任务中,数据集采用HCInew、HCIold、30Scenes、Occlusion和Reflective,评估了由2像素×2像素稀疏采样光场图像重建得到7像素×7像素光场图像的合成质量。

表2列出了本文模型与现有7种LFASR模型的量化对比结果。在HCInew和HCIold合成数据集上,本文模型在各项指标上均显著优于其他模型[26-29]。具体而言,在HCInew数据集上,本文模型的PSNR比次优的HLFSR-ASR模型高出1.28 dB,在HCIold数据集上也高出0.74 dB。在真实场景数据集30Scenes上,本文模型在PSNR指标上取得最高分,SSIM指标略低于最高分;在Occlusion和Reflective数据集上,本文模型也在SSIM指标方面表现优异。

2.4 LFSSR任务

在单一的光场空间超分辨率任务中,本文在

表 1　本文模型与 7 种模型在空间 (×2)-角度 (2×2→5×5) 同时超分辨率任务上的性能对比

Table 1　Performance comparisons of this paper with seven models on the joint spatial(×2) and angular(2×2→5×5) super-resolution task

LFASR模型	LFSSR模型	PSNR/dB, SSIM					
		30Scenes	Occlusion	Reflective	HCInew	HCIold	EPFL
DistgASR[16]	DistgSSR[16]	41.93 / 0.992 0	38.36 / 0.985 4	38.94 / 0.977 7	34.74 / 0.961 7	41.56 / 0.990 5	32.90 / 0.969 5
DistgASR[16]	EPIT[25]	41.85 / 0.991 8	38.27 / 0.985 1	38.89 / 0.976 8	35.21 / **0.965 3**	42.12 / 0.991 4	33.21 / 0.970 3
DistgASR[16]	MLFSR[14]	41.99 / 0.992 1	38.41 / 0.985 7	39.01 / 0.977 2	34.87 / 0.962 8	41.98 / 0.991 2	33.37 / 0.970 8
LF-EASR[7]	DistgSSR[16]	41.86 / 0.991 9	38.21 / 0.985 1	38.94 / 0.977 6	34.66 / 0.961 2	41.67 / 0.990 4	33.17 / 0.969 7
LF-EASR[7]	EPIT[25]	41.80 / 0.991 7	38.15 / 0.984 9	38.93 / 0.977 4	35.06 / 0.964 4	42.05 / 0.991 1	33.32 / 0.969 5
LF-EASR[7]	MLFSR[14]	41.77 / 0.991 7	38.12 / 0.984 9	38.91 / 0.977 2	34.67 / 0.961 3	41.68 / 0.990 7	33.28 / 0.970 0
LFSR[19]	LFSR[19]	41.89 / 0.991 9	38.30 / 0.985 3	38.86 / 0.977 2	34.46 / 0.959 9	41.72 / 0.990 7	32.98 / 0.969 2
本文模型	**本文模型**	**42.15 / 0.992 3**	**38.64 / 0.986 2**	**39.22 / 0.978 2**	**35.35 / 0.965 3**	**42.81 / 0.992 5**	**33.67 / 0.971 9**

注: 指标最优和次优分别作加粗和下划线处理。

HCInew、HCIold、EPFL、INRIA 和 STFgantry 数据集上与现有 9 种模型进行比对, 评估了光场图像在空间上放大 2 倍和 4 倍的合成质量, 表 3 列出了量化对比结果[30-34]。对比结果表明, 本文模型在 HCInew、HCIold、EPFL 和 INRIA 数据集上均取得最优的性能, 仅在 STFgantry 数据集上略低于最优性能。在×2 超分辨率实验中, 本文模型在 EPFL 数据集上的 PSNR 指标较次优的 HLFSR 模型提升了 0.39 dB。

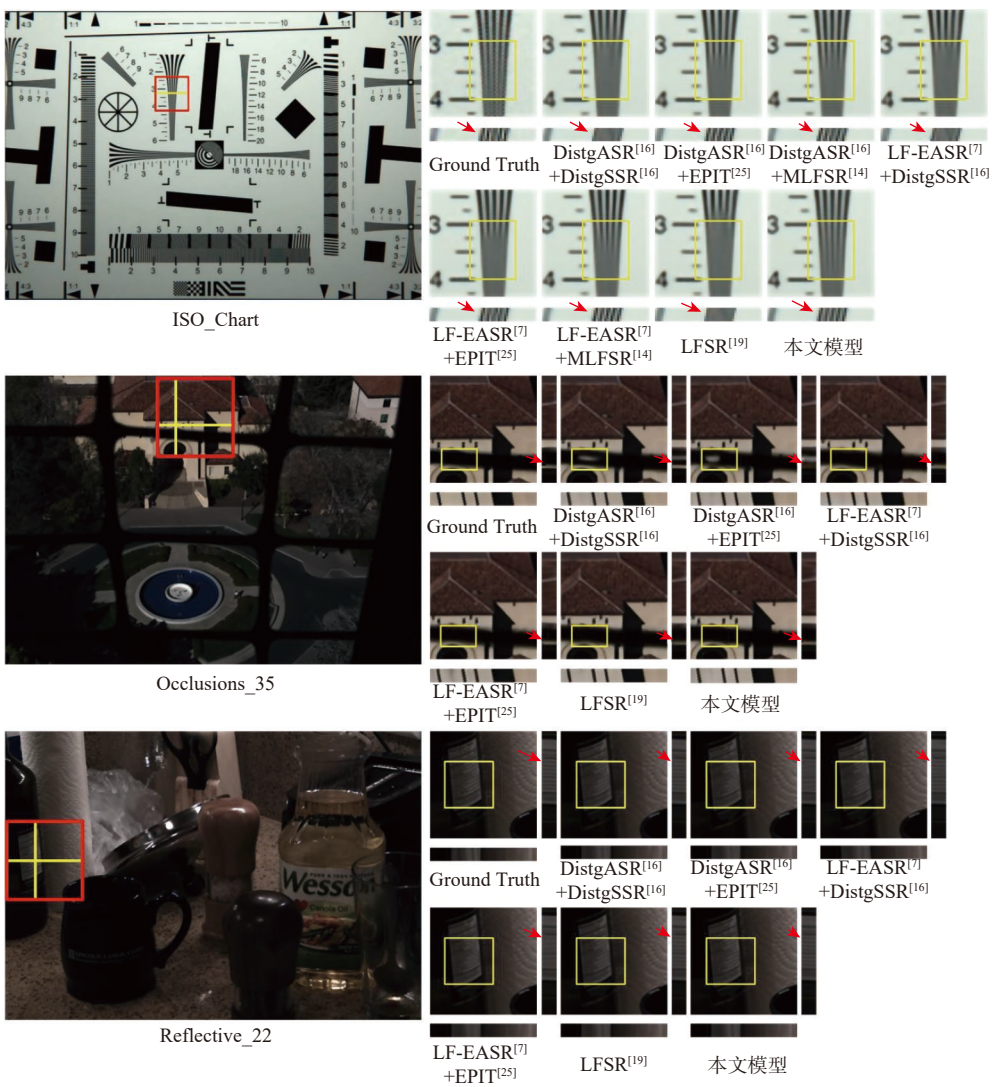


图 5　空间 (×2)-角度 (2×2→5×5) 同时超分辨率任务下不同模型生成的视觉效果对比

Fig. 5　Visual comparisons of different models on the joint spatial(×2) and angular(2×2→5×5) super-resolution task

表2 本文与7种模型在LFASR（2×2→7×7）任务上的PSNR和SSIM对比

Table 2 Performance comparisons on PSNR and SSIM of this paper with seven models in the LFASR task(2×2→7×7)

模型	发表年份	PSNR/dB, SSIM				
		HCInew	HCIold	30Scenes	Occlusion	Reflective
LF-EASR[7]	2022	35.86 / 0.975	41.54 / 0.971	43.44 / 0.995	39.80 / **0.992**	39.35 / 0.981
DistgASR[16]	2022	34.70 / 0.974	42.18 / 0.978	43.67 / 0.995	39.46 / 0.991	39.11 / 0.978
HLFSR-ASR[17]	2023	36.43 / 0.982	43.25 / 0.990	44.03 / **0.996**	39.87 / **0.992**	**39.60** / 0.981
Liu等[26]	2023	35.29 / 0.954	41.92 / 0.969	43.33 / 0.988	39.67 / 0.983	39.20 / 0.963
Li和Zhong[27]	2024	35.21 / 0.975	42.36 / 0.979	43.95 / **0.996**	39.68 / 0.991	39.36 / 0.981
Liu等[28]	2024	35.41 / 0.954	42.02 / 0.969	43.38 / 0.995	39.71 / 0.991	39.31 / 0.982
ViewFormer[29]	2025	35.37 / 0.977	41.93 / 0.975	43.82 / 0.995	39.86 / **0.992**	39.57 / 0.981
本文		**37.71** / **0.983**	**43.99** / **0.993**	**44.15** / 0.995	**40.07** / **0.992**	39.45 / **0.984**

注: 指标最优和次优分别作加粗和下划线处理。

表3 本文与9种模型在LFSSR任务(放大2倍和4倍)上的PSNR和SSIM对比

Table 3 Performance comparisons on PSNR and SSIM of this paper with nine models in the LFSSR tasks(×2 and ×4)

模型	发表年份	PSNR/dB, SSIM									
		EPFL		HCInew		HCIold		INRIA		STFgantry	
		放大2倍	放大4倍	放大2倍	放大4倍	放大2倍	放大4倍	放大2倍	放大4倍	放大2倍	放大4倍
DPT[30]	2022	34.49 0.975 8	28.94 0.917 0	37.36 0.977 1	31.20 0.918 8	44.30 0.994 3	37.41 0.972 1	36.41 0.984 3	30.96 0.952 3	39.43 0.992 6	31.15 0.948 8
DistgSSR[16]	2022	34.81 0.978 7	28.99 0.919 5	37.96 0.979 6	31.38 0.921 7	44.94 0.994 9	37.56 0.973 2	36.59 0.985 9	30.99 0.951 9	40.40 0.994 2	31.65 0.953 5
EPIT[25]	2023	34.83 0.977 5	29.34 0.919 7	38.23 0.981 0	31.51 0.923 1	45.08 0.994 9	37.68 0.973 7	36.67 0.985 3	31.37 0.952 6	**42.17** **0.995 7**	**32.18** 0.957 1
HLFSR-SSR[17]	2023	35.31 0.980 0	29.20 0.922 2	38.32 0.980 7	31.57 0.923 8	44.98 0.995 0	37.78 0.972 4	37.06 0.986 7	31.24 0.953 4	40.85 0.994 7	31.64 0.953 7
LF-DET[31]	2023	35.26 0.979 7	**29.47** 0.923 0	38.31 0.980 7	31.56 0.923 5	44.99 0.995 0	37.84 0.974 4	36.95 0.986 4	31.39 0.953 4	41.76 0.995 5	32.14 **0.957 3**
徐欣宜等[32]	2023	34.49 0.976 7	28.70 0.914 8	37.47 0.979 0	31.13 0.920 8	44.25 0.994 3	37.21 0.972 4	36.37 0.984 3	30.73 0.950 3	38.89 0.991 6	30.61 0.943 2
MLFSR[14]	2024	35.22 0.980 1	29.28 0.921 8	38.14 0.980 3	31.56 0.923 5	44.90 0.995 0	37.83 0.974 5	36.92 0.986 5	31.24 0.953 1	40.98 0.994 9	32.03 0.956 7
LF-TGUnet[33]	2024	34.67 0.977 6	28.96 0.918 9	37.80 0.979 1	31.42 0.921 5	44.78 0.994 8	37.63 0.973 4	36.46 0.985 7	31.28 0.952 9	39.95 0.993 6	31.36 0.951 1
LF-GIANet[34]	2025	34.56 0.978 2	29.14 0.919 8	38.01 0.979 8	31.40 0.921 8	44.80 0.994 8	37.60 0.973 3	36.45 0.985 8	31.25 0.952 2	40.45 0.994 5	31.30 0.950 5
本文		**35.70** **0.980 9**	**29.47** **0.924 4**	**38.49** **0.981 5**	**31.61** **0.924 4**	**45.31** **0.995 3**	**37.86** **0.974 8**	**37.28 /** **0.987 2**	**31.45** **0.955 0**	41.45 0.995 3	31.72 0.955 3

注: 指标最优和次优分别作加粗和下划线处理。

2.5 低空场景下LFASSR泛化性能

为评估本文模型在低空场景下应用的泛化性能，从UrbanLF-Real[24]数据集中选取12个典型低空应用的光场图像，包含建筑群、道路交通及桥梁结构等多种典型城市低空场景，如图4所示，能有效模拟无人机、城市监控等低空视角下的光场成像条件。

在泛化实验中，直接采用2.2节中已训练的LFASSR模型，未进行任何额外的微调或重新训练。表4列出了全部12个低空场景光场测试图像的超分辨率结果。尽管本文模型未在该数据集上进行专门训练，但仍在PSNR和SSIM指标上取得优异性能，并在所有比对方法中获得了最高的平均指标值。图6展示了其中2个低空场景光场测试图像的超分辨率重建结果，在红色框选区域中，对比超分辨率前后的局部细节，可以明显看出，对应区域的结构与纹理信息得到了显著恢复。图7选择低空光场测试集中的Image128场景（即图6中的

第 2 个场景), 给出本文方法与其他 3 种模型在该场景下 LFASSR 的指标分布情况。可以看出, 本文方法在 PSNR 和 SSIM 这 2 项指标上不仅取得最高的平均值, 其分布方差也最小, 表明本文方法在整体重建质量和跨视角稳定性方面均具有明显优势。以上泛化实验表明, 本文方法在复杂城市低空场景中仍能有效恢复光场的空间与角度信息, 说明本文模型具有良好的泛化能力。

表 4　低空场景光场图像的空间 (×2)-角度 (2×2→5×5) 同时超分辨率性能对比

Table 4　Performance comparisons on the joint spatial (×2) and angular (2×2→5×5) super-resolution for low-altitude light field images

场景编号	PSNR/dB, SSIM							
	DistgASR[16] +DistgSSR[16]	DistgASR[16] +EPIT[25]	DistgASR[16] +MLFSR[14]	LF-EASR[7] +DistgSSR[16]	LF-EASR[7] +EPIT[25]	LF-EASR[7] + MLFSR[14]	LFSR[19]	本文模型
Image36	37.09/0.981 7	37.06/0.981 6	37.15/0.981 9	36.85/0.980 4	37.01/0.981 3	37.00/0.980 9	36.82/0.980 4	**37.20/0.982 3**
Image62	36.34/0.978 5	36.41/0.979 1	**36.64/0.979 3**	36.07/0.977 7	36.42/0.978 8	36.34/0.978 2	36.04/0.977 5	**36.64/0.979 3**
Image63	36.08/0.983 4	36.25/0.984 0	36.30/0.984 1	36.17/0.983 9	36.19/0.983 8	36.13/0.983 6	35.89/0.982 8	**36.44/0.984 5**
Image81	37.87/0.980 7	37.97/0.981 2	37.92/0.980 9	37.95/0.981 2	37.87/0.980 8	37.85/0.980 6	37.68/0.980 0	**38.04/0.981 4**
Image82	36.87/0.978 7	37.01/0.979 3	37.03/0.979 2	36.68/0.977 7	36.93/0.978 4	36.71/0.977 8	36.50/0.977 1	**37.30/0.980 5**
Image85	37.56/0.986 1	37.79/0.986 5	37.74/0.986 3	37.41/0.985 7	37.76/0.986 0	37.42/0.985 5	37.10/0.984 7	**38.10/0.987 2**
Image86	39.56/0.978 9	39.57/0.978 9	39.48/0.978 6	39.59/0.979 0	39.52/0.978 7	39.55/0.978 8	39.46/0.978 5	**39.64/0.979 2**
Image89	37.26/0.978 9	37.23/0.979 0	37.33/0.979 2	37.20/0.978 7	37.26/0.978 9	37.23/0.978 7	37.03/0.978 2	**37.37/0.979 5**
Image98	37.21/0.983 7	**37.41/0.984 4**	37.29/0.983 9	37.06/0.983 1	37.29/0.983 8	37.11/0.983 1	37.07/0.983 2	37.36/0.984 3
Image128	39.52/0.991 0	39.69/0.992 0	39.56/0.992 1	39.51/0.991 9	39.69/0.991 8	39.53/0.991 6	39.14/0.990 1	**39.99/0.992 8**
Image187	37.23/0.991 4	37.38/0.992 4	37.21/0.992 0	37.48/0.992 1	37.25/0.991 6	37.10/0.991 4	36.80/0.991 0	**37.69/0.992 6**
Image264	35.68/0.981 2	36.04/0.982 6	**36.19/0.983 0**	34.69/0.977 5	36.02/0.981 9	34.62/0.976 3	35.34/0.979 8	36.01/0.982 9

注: 指标最优和次优分别作加粗和下划线处理; 从左到右8个模型的平均值分别为37.36/0.982 9、37.48/0.983 4、37.49/0.983 4、37.22/0.982 4、37.43/0.983 0、37.22/0.982 2、37.07/0.982 0、37.65/0.983 9。



图 6　低空场景光场图像的空间 (×2)-角度 (2×2→5×5) 同时超分辨率重建视觉结果

Fig. 6　Visual results of joint spatial (×2) and angular (2×2→5×5) super-resolution on low-altitude light field images task (the second scene in Fig.6)

2.6　消融实验

为验证本文方法和策略的有效性, 分别在 LFASR、LFSSR、LFASSR 任务上进行消融实验。

表 5 评估了 4D 光场高斯泼溅模块在 LFASR 和 LFSSR 任务中的贡献。在 w/o 4DLFGS 情况下, 采用现有局部隐式图像函数(local implicit image function, LIIF)方法[35] 替换 4D 光场高斯泼溅模块, 各项性能指标都出现不同程度的退化, 尤其在 HCInew 数据集上, PSNR 指标出现明显下降, 表明在 EPI 特征图上进行高斯泼溅建模和渲染能有效提升超分辨率重建质量。

表 6 在 LFASSR 任务下评估了交叉注意力模块、自注意力模块和 4DLFGS 的有效性, 共在 6 个数据集上进行实验。比对结果表明, 无论是真实场景还是合成场景, 去除任意一个模块都会导致模型性能指标下降, 进一步证明本文模块和策略的有效

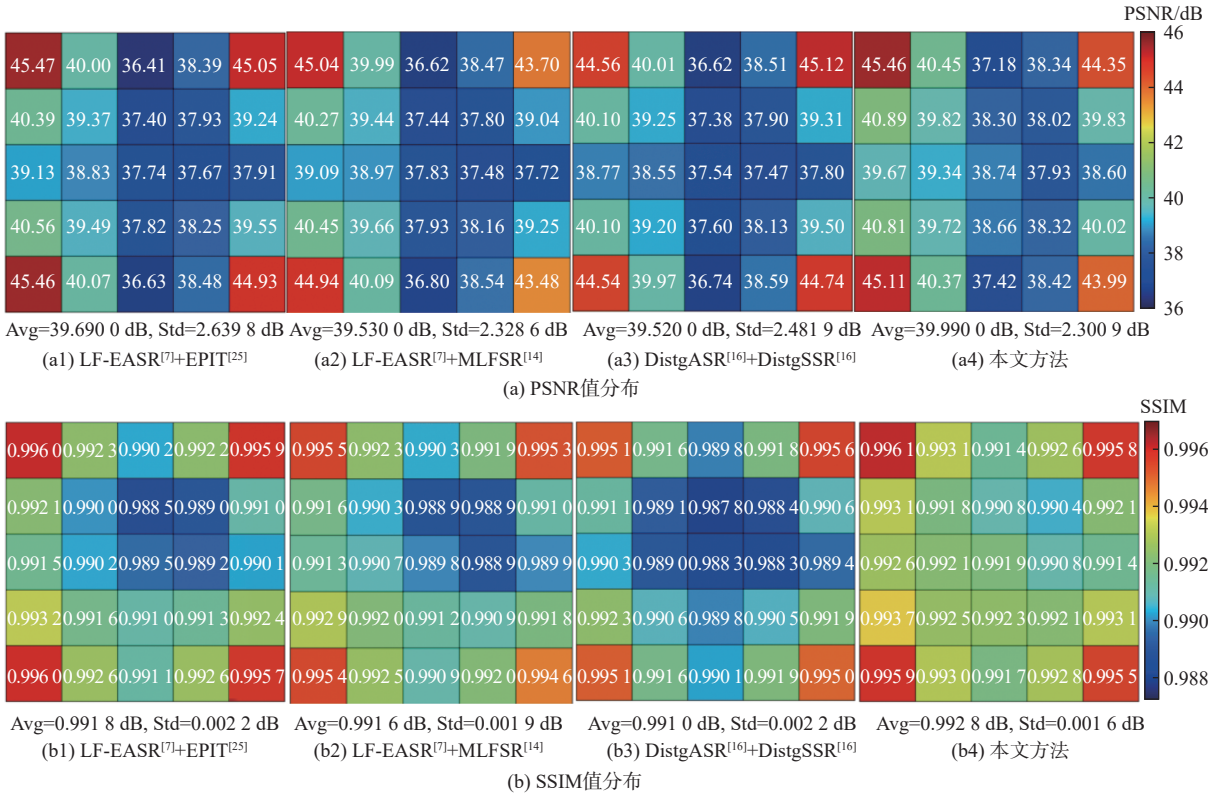


图 7 空间 (×2)-角度 (2×2→5×5) 同时超分辨率任务下 PSNR 和 SSIM 值分布的可视化比对 (图 6 中第 2 个场景)

Fig. 7 Visual comparison of PSNR and SSIM distributions under joint spatial (×2) and angular (2×2→5×5) super-resolution

表 5 LFASR 和 LFSSR 消融实验

Table 5 Ablation study on the LFASR and LFSSR tasks

超分辨率任务类型	数据集	PSNR (dB)/SSIM	
		w/o 4DLFGS	本文方法
LFASR	HCInew	34.27/0.970 0	37.71/0.982 9
	HCIold	42.27/0.979 0	43.99/0.993 1
LFSSR	HCInew	38.32/0.981 0	38.49/0.981 5
	HCIold	45.01/0.995 0	45.31/0.995 3

表 6 LFASSR 任务消融实验

Table 6 Ablation study on the LFASSR task

交叉注意力模块	自注意力模块	4DLFGS	PSNR/dB, SSIM						参数量	浮点运算速度/($10^9 \cdot s^{-1}$)
			HCInew	HCIold	EPFL	30Scenes	Occlusions	Reflective		
√			34.55/0.960 4	41.88/0.991 2	33.22/0.970 2	41.84/0.991 8	38.28/0.985 3	38.96/0.977 3	1.483×10^7	76.12
	√		34.41/0.959 6	41.70/0.991 2	33.26/0.969 9	41.90/0.991 9	38.34/0.985 5	39.06/0.978 1	1.485×10^7	76.30
√	√		35.12/0.963 9	42.40/0.992 0	33.51/0.971 1	41.96/0.992 0	38.45/0.985 7	39.12/0.978 6	1.638×10^7	85.63
	√	√	35.09/0.963 9	42.42/0.992 0	33.51/0.971 5	42.11/0.992 2	38.53/0.985 8	39.13/0.978 0	1.525×10^7	119.93
√	√	√	35.35/0.965 3	42.81/0.992 5	33.67/0.971 9	42.15/0.992 3	38.64/0.986 2	39.22/0.978 2	1.678×10^7	129.26

性。另外,表中还给出了各种模型的参数量和浮点运算速度,可以看出,本文提出的 4 DLFGS 参数量仅高出 0.4×10^5。

3 结 论

本文实验结果可得出以下结论:

1) 本文方法能够有效实现光场空间与角度同时超分辨率。在空间分辨率放大 2 倍、角度分辨率由 (2×2) 提升至 (5×5) 的联合超分辨率任务中,本文方法在 30Scenes、Occlusion、Reflective、HCInew、HCIold 和 EPFL 6 个数据集上的 PSNR 均取得最优结果。表明同时进行空间和角度重建能够获得稳定的性能优势。

2) 本文方法不仅适用于空间-角度同时超分辨

率, 在单一角度或空间超分辨率任务中也具有较好的适用性。其中, 在 (2×2->7×7) 角度超分辨率任务的 HCInew 和 HCIold 数据集上, PSNR 分别达到 37.71 dB 和 43.99 dB, 较次优方法提高 1.28 dB 和 0.74 dB; 在空间超分辨率任务中, 除 STFgantry 数据集外, 在 HCInew、HCIold、EPFL 和 INRIA 数据集上均取得最优或具有竞争力的重建性能, 说明该模型对不同光场超分辨率任务具有较好的通用性。

3) 本文方法对未参与训练的低空场景具有较好的泛化能力。在不针对 UrbanLF-Real 数据集进行额外训练或微调的情况下, 12 个建筑、道路、交通和桥梁等典型低空场景上的平均 PSNR 和 SSIM 分别达到 37.65 dB 和 0.983 9, 均优于其他对比方法; 在 Image128 场景中, 本文方法的平均 PSNR 和 SSIM 分别达到 39.99 dB 和 0.992 8, 同时两项指标的跨视角分布方差均为对比方法中最小, 说明其在真实低空场景下具有较好的重建质量和跨视角稳定性。

4) 消融实验表明, 各模块对超分辨率性能均具有实际贡献, 且 4D 光场高斯泼溅对角度信息重建的作用尤为明显。在空间-角度同时超分辨率任务中, 去除交叉注意力、自注意力或 4D 光场高斯泼溅模块均会造成 6 个数据集上的性能下降; 在角度超分辨率任务中, 以 LIIF 替代 4D 光场高斯泼溅后, HCInew 数据集上的 PSNR 由 37.71 dB 降至 34.27 dB, 下降 3.44 dB, 说明基于 EPI 的高斯泼溅建模能够有效增强光场角度域几何信息的恢复能力。

5) 本文方法在极端光照、逆光和低光照条件下仍可能受到成像噪声及细节退化的影响, 对高频纹理和跨视角一致性的恢复能力尚有提升空间; 同时, 现有模型主要针对固定放大倍率设计, 在更高倍率及任意尺度超分辨率任务中的计算效率和重建稳定性仍需进一步验证和优化。后续研究可重点面向复杂真实场景和任意尺度光场重建进一步改进模型。

参考文献 (References)

[1] Ng R, Levoy M, Brédif M, et al. Light field photography with a hand-held plenoptic camera[D]. Palo Alto: Stanford University, 2005: 5-11.

[2] Lyu W Q, Sheng H, Ke W, et al. Advances in light field spatial super-resolution: a comprehensive literature survey[J]. IEEE Access, 2025, 13: 18470-18497.

[3] Hu 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.

[4] Kalantari N K, Wang T C, Ramamoorthi R. Learning-based view synthesis for light field cameras[J]. ACM Transactions on Graphics, 2016, 35(6): 1-10.

[5] Zhang S, Chang S, Shen Z Q, et al. Micro-lens image stack upsampling for densely-sampled light field reconstruction[J]. IEEE Transactions on Computational Imaging, 2021, 7: 799-811.

[6] Jin J, Hou J H, Yuan H, et al. Learning light field angular super-resolution via a geometry-aware network[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 11141-11148.

[7] Liu G S, Yue H J, Wu J M, et al. Efficient light field angular super-resolution with sub-aperture feature learning and macro-pixel upsampling[J]. IEEE Transactions on Multimedia, 2023, 25: 6588-6600.

[8] Wang L Y, Ren L F, Wei X Y, et al. Light field angular super-resolution based on intrinsic and geometric information[J]. Knowledge-Based Systems, 2023, 270: 110553.

[9] Mao Y F, Xiao Z Y, An P, et al. Deep sparse-to-dense inbetweening for multi-view light fields[J]. IEEE Transactions on Image Processing, 2025, 34: 6302-6317.

[10] Yuan Y, Cao Z Q, Su L J. Light-field image superresolution using a combined deep CNN based on EPI[J]. IEEE Signal Processing Letters, 2018, 25(9): 1359-1363.

[11] Wang Y L, Liu F, Zhang K B, et al. LFNet: a novel bidirectional recurrent convolutional neural network for light-field image super-resolution[J]. IEEE Transactions on Image Processing, 2018, 27(9): 4274-4286.

[12] Zhang S, Lin Y F, Sheng H. Residual networks for light field image super-resolution[C]// Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 11038-11047.

[13] Liang Z Y, Wang Y Q, Wang L G, et al. Light field image super-resolution with transformers[J]. IEEE Signal Processing Letters, 2022, 29: 563-567.

[14] Gao R S, Xiao Z Y, Xiong Z W. Mamba-based light field super-resolution withEfficient subspace scanning[C]//Proceedings of the Computer Vision–ACCV 2024. Berlin: Springer, 2024: 421-437.

[15] Chen S R, Chen L, Wu D F, et al. Enhancing light field image super-resolution through Mamba-based spatial–angular correlation learning[J]. The Visual Computer, 2025, 41(15): 12649-12662.

[16] Wang Y Q, Wang L G, Wu G C, et al. Disentangling light fields for super-resolution and disparity estimation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(1): 425-443.

[17] Van Duong V, Huu T N, Yim J, et al. Light field image super-resolution network via joint spatial-angular and epipolar information[J]. IEEE Transactions on Computational Imaging, 2023, 9: 350-366.

[18] Yoon Y, Jeon H G, Yoo D, et al. Learning a deep convolutional network for light-field image super-resolution[C]// Proceedings of the 2015 IEEE International Conference on Computer Vision Workshop. Piscataway: IEEE Press, 2016: 57-65.

[19] Liu G S, Yue H J, Li K, et al. Adaptive pixel aggregation for joint spatial and angular super-resolution of light field images[J]. Information Fusion, 2024, 104: 102183.

[20] Liu D Y, Li S Z, Mao Y F, et al. Learning implicit and detail-enhanced network for light field image spatial-angular super-resolution[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2026, 36(2): 1544-1557.

[21] Wu T, Yuan Y J, Zhang L X, et al. Recent advances in 3D Gaussian splatting[J]. Computational Visual Media, 2024, 10(4): 613-642.

[22] Hu J T, Xia B, Chen B, et al. GaussianSR: high fidelity 2D Gaussian splatting for arbitrary-scale image super-resolution[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2025, 39(4): 3554-3562.

[23] Zhang H, Zhou W H, Lin L L, et al. Cascade residual learning based adaptive feature aggregation for light field super-resolution[J]. Pattern Recognition, 2025, 165: 111616.

[24] Sheng H, Cong R X, Yang D, et al. UrbanLF: a comprehensive light field dataset for semantic segmentation of urban scenes[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(11): 7880-7893.

[25] Liang Z Y, Wang Y Q, Wang L G, et al. Learning non-local spatial-angular correlation for light field image super-resolution[C]// Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2023: 12342-12352.

[26] Liu D Y, Mao Y F, Zhou X F, et al. Learning a multilevel cooperative view reconstruction network for light field angular super-resolution[C]// Proceedings of the 2023 IEEE International Conference on Multimedia and Expo. Piscataway: IEEE Press, 2023: 1271-1276.

[27] Li D C, Zhong R. Light field image angular super-resolution using edge features[C]//Proceedings of the 2024 6th International Conference on Robotics and Computer Vision. Piscataway: IEEE Press, 2024: 132-138.

[28] Liu D Y, Mao Y F, Zuo Y F, et al. Light field angular super-resolution network based on convolutional transformer and deep deblurring[J]. IEEE Transactions on Computational Imaging, 2024, 10: 1736-1748.

[29] Wang S Z, Lu Y, Xia W, et al. Light field angular super-resolution by view-specific queries[J]. The Visual Computer, 2025, 41(5): 3565-3580.

[30] Wang S Z, Zhou T F, Lu Y, et al. Detail-preserving transformer for light field image super-resolution[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(3): 2522-2530.

[31] Cong R X, Sheng H, Yang D, et al. Exploiting spatial and angular correlations with deep efficient transformers for light field image super-resolution[J]. IEEE Transactions on Multimedia, 2023, 26: 1421-1435.

[32] 徐欣宜, 邓慧萍, 向森, 等. 基于特征交互融合与注意力的光场图像超分辨率[J]. 激光与光电子学进展, 2023, 60(14): 1410017.
Xu X Y, Deng H P, Xiang S, et al. Light field image super-resolution based on feature interaction fusion and attention mechanism[J]. Laser & Optoelectronics Progress, 2023, 60(14): 1410017(in Chinese).

[33] 黄莉, 吕天琪, 武迎春, 等. 基于双路引导更新的光场图像超分网络[J]. 光电工程, 2024, 51(12): 240222.
Huang L, Lv T Q, Wu Y C, et al. Two-way guided updating network for light field image super-resolution[J]. Opto-Electronic Engineering, 2024, 51(12): 240222(in Chinese).

[34] Zhang W, Li H J, Ke W. LF-GIANet: cascaded global-view information adaptation-guided network for light field image super-resolution[J]. Multimedia Systems, 2025, 31(2): 73.

[35] Chen Y B, Liu S F, Wang X L. Learning continuous image representation with local implicit image function[C]// Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 8624-8634.

Light field gaussian splatting representation and joint spatial-angular super-resolution for low-altitude scenarios

MENG Jiahan[1]，YANG Chi[1]，LI Chuanjun[2]，LIN Lili[2]，ZHOU Wenhui[1,*]

(1. School of Computer Science and Technology，Hangzhou Dianzi University，Hangzhou 310018，China；
2. School of Information and Electronic Engineering，Zhejiang Gongshang University，Hangzhou 310018，China)

Abstract: Unmanned aerial vehicle (UAV) perception and environmental monitoring are two low-altitude commercial applications where light field imaging technology shows significant promise by simultaneously capturing the spatial and angular information of light rays. However, existing light field cameras and super-resolution methods struggle to jointly optimize spatial and angular resolutions, limiting their effectiveness in tasks requiring precise perception and high-fidelity reconstruction. To address this challenge, this paper proposes a light field Gaussian splatting representation and a light field angular spatial super-resolution (LFASSR) method tailored for low-altitude scenarios. Specifically, we extend the conventional 3D Gaussian splatting framework into the 4D light field domain and propose a 4D light field Gaussian splatting representation. Leveraging the geometric structure of light field epipolar plane images (EPIs), we decompose the 4D representation into cascaded 2D Gaussian splatting models along horizontal and vertical directions, enabling more efficient modeling of light field geometry. Based on this representation, we develop a unified light field super-resolution framework. The framework first extracts and fuses features from the spatial, angular, and EPI domains, then constructs directional geometric feature spaces for horizontal and vertical views, where 2D Gaussian splatting is applied separately. Finally, high-quality dense-view light field images with enhanced spatial resolution are rendered via EPI-based synthesis. The suggested methodology greatly outperforms state-of-the-art methods in joint spatial-angular super-resolution, as shown by extensive experimentation on public light field benchmarks and a recently gathered low-altitude light field dataset. It achieves the best performance in terms of peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM), while also producing visually superior results with sharp detail recovery and smooth, natural disparity transitions. Moreover, the method exhibits strong generalization capability and competitive performance in standalone light field spatial super-resolution (LFSSR) and light field angular super-resolution (LFASR) tasks.

Keywords: light field; light field super-resolution; light field Gaussian splatting; spatial-angular joint super-resolution; epipolar plane geometry

Received：2025-11-17；**Accepted**：2025-12-12；**Published Online**：2026-02-02 11：35

URL：link.cnki.net/urlid/11.2625.V.20260130.1851.004

Foundation items：National Natural Science Foundation of China (62471437); ‘‘Pioneer’’ and ‘‘Leading Goose’’ Research and Development Program of Zhejiang (2023C01030)

* **Corresponding author**. E-mail：zhouwenhui@hdu.edu.cn