

文章编号: 1671-7449(2026)03-0335-09

J. Test Measur. Technol., 2026, 40(3): 335-343.

一种高效的光场焦点堆栈的超分辨率网络

牛晶慧¹, 张峻彬², 袁仲云^{1*}, 程永强², 赵 纯³

(1. 太原理工大学 集成电路学院, 山西 太原 030024; 2. 太原理工大学 电子信息工程学院, 山西 太原 030024;
3. 小米科技有限责任公司, 北京 100084)

摘要: 光场相机从多个视角捕捉场景, 提供全面的图像信息。然而, 由此产生的数据存在低图像质量问题, 并且当前研究普遍存在多聚焦特征开发不足的局限, 对深度特征能力的开发不足直接制约了光场增强的视觉效果。鉴于此, 利用深度信息将子孔径图像转换为焦点堆栈并对其进行解析, 利用焦点堆栈具有不同景深的特点, 通过设计滤波器得到焦点堆栈所对应的深度图, 将图像分类为聚焦或非聚焦, 然后选择性地应用超分辨率网络来增强聚焦图像, 该网络采用维度拉伸策略将低分辨率焦点堆栈和退化参数对应的退化图作为卷积神经网络输入。在公开和自建光场数据集上进行了实验, 结果表明所提方法的图像信息熵较LF-IINet方法平均提升前景0.21%, 平均提升后景0.22%, 较Distg-SSR方法平均提升前景0.21%, 平均提升后景0.28%; 在图像高频能量方面, 所提方法较LF-IINet方法平均提升前景10.78%, 平均提升后景10.56%, 较Distg-SSR方法平均提升前景11.33%, 平均提升后景11.02%。通过深度引导的聚焦增强与非聚焦模糊策略实现了细节增强与噪声抑制的平衡, 为光场成像在计算摄影领域的应用提供了高效可靠的解决方案。

关键词: 光场; 多聚焦特性; 焦点堆栈; 深度图; 超分辨率

中图分类号: TP311 文献标识码: A DOI: 10.62756/csjs.1671-7449.2026025

引用格式: 牛晶慧, 张峻彬, 袁仲云, 等. 一种高效的光场焦点堆栈的超分辨率网络[J]. 测试技术学报, 2026, 40(3): 335-343.

NIU Jinghui, ZHANG Junbin, YUAN Zhongyun, et al. An efficient super-resolution network for light field focal stack[J]. Journal of Test and Measurement Technology, 2026, 40(3): 335-343.

An Efficient Super-Resolution Network for Light Field Focal Stack

NIU Jinghui¹, ZHANG Junbin², YUAN Zhongyun^{1*}, CHENG Yongqiang², ZHAO Chun³

(1. College of Integrated Circuits, Taiyuan University of Technology, Taiyuan 030024, China;
2. College of Electronic Information Engineering, Taiyuan University of Technology, Taiyuan 030024, China;
3. Xiaomi Corporation, Beijing 100084, China)

Abstract: Light field (LF) cameras capture scenes from multiple perspectives, providing comprehensive image information. However, the resulting data often suffers from low image quality, and current research generally exhibits limitations in underutilizing multi-focus characteristics. The insufficient exploitation of deep feature capabilities directly constrains the visual enhancement effects of LF imaging. To address this issues, converting sub-aperture images (SAIs) is proposed into a focal stack (FS) using depth information and analyzing them. Leveraging the varying depth-of-field characteristics of FS, depth maps are generated through specifically designed filters to classify images as focused or defocused. A selective super-resolution (SR) network is then

【收稿日期】2025-05-04

【基金项目】国家自然科学基金资助项目(52275568); 山西省重点研发计划资助项目(202102150401011)

【作者简介】牛晶慧(1999—), 女, 硕士生, 主要从事光场图像处理研究。E-mail: 19834503281@163.com。

*【通信作者】袁仲云(1982—), 男, 教授, 博士, 主要从事光场图像技术研究。E-mail: yuanzhongyun@tyut.edu.cn。

applied to enhance focused images, where a dimensionality stretching strategy is employed to integrate low-resolution FS with degradation maps corresponding to deterioration parameters as input to convolutional neural networks. Experiments were conducted on both public and self-built light field datasets. The results show that the proposed method has an average increase of 0.21% in the image information entropy for the foreground and 0.22% for the background compared to the LF-IINet method, and an average increase of 0.21% for the foreground and 0.28% for the background compared to the Distg-SSR method. In terms of the high-frequency energy of the image, the proposed method has an average increase of 10.78% for the foreground and 10.56% for the background compared to the LF-IINet method, and an average increase of 11.33% for the foreground and 11.02% for the background compared to the Distg-SSR method. Through depth-guided focused enhancement and defocused blurring strategies, this approach achieves a balance between detail enhancement and noise suppression. The framework provides an efficient and reliable solution for advancing LF imaging applications in computational photography.

Key words: light field (LF); multi-focus characteristics; focal stack; depth maps; super-resolution

0 引言

随着光场成像技术^[1]的快速发展,目前已经引入了多种设计来捕捉光场数据^[2-4]。光场成像从场景的所有方向捕捉光信息,生成包含颜色、几何结构、方向和距离信息的4D数据,已被应用在多个方面,例如拍摄后重新聚焦^[5-6]、深度感知^[7-9]、反射率估计^[10-13]、3D重建^[14]、显著性检测^[15]和运动去模糊^[16]等。关于光场焦点堆栈,全聚焦图像和深度图有助于减轻与大数据量相关的高计算成本^[17]。Takahashi等^[18]开发了一种基于焦点堆栈的光场显示技术,该技术允许同时从多个角度进行即时观看,且不牺牲每个观看方向的分辨率。总体而言,目前研究表明,光场是虚拟现实和增强现实的一种有前景的媒介。传统成像机制仅能获取三维场景在二维平面的投影映射,不可避免地损失了原始场景的几何信息。相较之下,光场相机的成像技术可通过单次曝光便可实现多角度和不同景深记录,不仅可实现空间-角度双维度信息耦合,更具备从全光函数中解析任意方向辐射分量的能力,突破了过往成像系统信息缺失的物理限制。因此,利用光场相机可以记录不同景深,能够为光场图像超分辨率研究提供数据支持。近年来,卷积神经网络在低频图像处理任务中得到了广泛研究。然而,由于现实场景中复杂的场景结构,在低分辨率图像中,空间和角度与变化的视差高度交织,给卷积神经网络有效利用信息线索带来了重大挑战。为了解决这一问题,现有方法通常利用冗余的光场结构(即通过记录不同角度视图中的相似图像内容)并通过对邻近视图应用卷积来降低

光场图像的维度^[19-20]、极平面图像^[21-24]或3D子孔径图像^[25-28]。然而,这些方法往往无法充分利用光场数据中固有的丰富深度信息。通过更有效地利用光场图像结构,可以显著提高基于卷积神经网络方法的性能。因此,本文提出了一种深度信息引导的光场超分辨率网络。区别于上述传统方法缺失对深度信息的充分利用,本文构建深度信息与卷积网络间的协同机制:利用深度信息将子孔径图像转换为焦点堆栈对其进行解析,利用焦点堆栈具有不同景深的特点,通过设计滤波器得到焦点堆栈所对应的深度图,将图像分类为聚焦或非聚焦,然后选择性地应用超分辨率网络来增强聚焦图像。

1 光场成像

传统相机中,光信息被记录为二维数据,仅能表示平面信息,其感光元件仅获取各像素单元的光谱特性和辐射强度。相比之下,全光相机记录的数据维度超越传统亮度与色度参数,更涵盖入射光线的空间传播矢量特征。所捕获的附加信息可用于光线位置的重建,更有助于三维立体图形的建构。

微透镜阵列成像技术的理论可追溯至Lippmann于1908年提出的“整体摄影”^[29],该体系采用窄基距球面微透镜阵列对场景进行多视角采样,生成具有视差特征的影像阵列。如图1所示的模型后来被概念化为商用光场相机^[30],其核心由三大部分构成:主透镜、透镜阵列和传感器。每个透镜形成透镜光圈的一个微型部分,用来测量光线在该透镜上的定向分布。

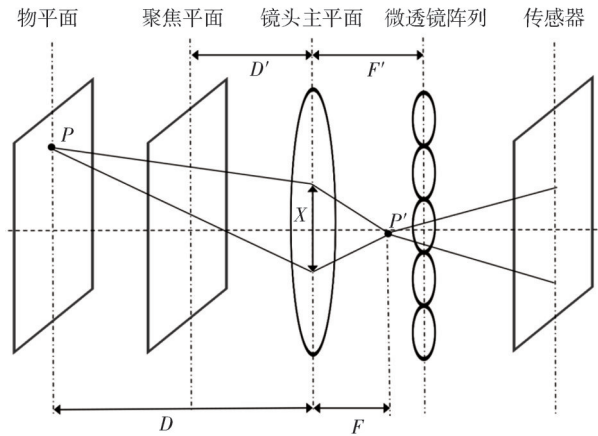


图 1 光场成像原理

Fig. 1 Principle of light field imaging

作为整体视图的原始数据看起来与传统图像相似,但在显微镜下,每个微透镜捕获的子孔径图像是清晰可见的。透镜阵列成像系统通过解析四维光场函数,物平面上的点 P 发出光线传播至聚焦平面形成初步会聚,随后穿过距离 D' 到达镜头主平面。此时,光线继续向微透镜阵列传播,每个微透镜(对应 X 长度范围)将来自物点 P 的不同角度光线分离,投射至特定位置形成像点 P' ,再经过微透镜阵列映射至传感器。在此过程中,参数 D' (聚焦平面-镜头主平面)与 F' (镜头主平面-微透镜阵列)协同调控光线角度分布与空间几何关系,而微透镜阵列通过物理编码将光线的四维信息(位置+方向)映射至传感器像素,形成包含全光场数据的二维记录。通过分析各个参数之间的关系可以计算出光场相机采集物

体的深度值 S 为视差, X 、 D' 和 F' 为相机固定参数,据此可计算深度 D 。

2 超分辨率网络

2.1 网络框架

使用一种高效的超分辨率网络进行光场焦点堆栈。图 2 描述了所提出的面向光场图像增强的超分辨率画质研究解决方案。原始焦点堆栈由 Lytro Illum 光场相机采集并解码获取,之后进一步生成光场子孔径图像,然后,将光场子孔径图像转换为焦点堆栈以提取深度信息,根据深度信息将焦点堆栈分为聚焦图和未聚焦图,并选择性地对聚焦图像进行超分辨率处理。超分辨率网络设计了一种高效且可扩展的卷积神经网络框架,可同时处理模糊核、噪声输入和低分辨率网络。利用不同焦平面图像的清晰度信息估计每个像素的深度,随后将焦栈分为聚焦与未聚焦图像,并仅对聚焦图像的前景和背景进行超分辨率处理。未聚焦图像可添加不同噪声水平的模糊或保持原状。模糊未聚焦图像有助于更突出聚焦图像的主体。因此,本文将拼接后尺寸为 $W \times H \times (C + t + 1)$ 的低分辨率焦点堆栈图像和退化图作为输入,对低分辨率图像进行超分辨率处理(缩放因子为 r)。采用一系列 3×3 卷积层进行非线性映射,除最后一层,每层卷积均采用“卷积+批归一化+激活函数”结构,最后一层仅包含卷积操作。

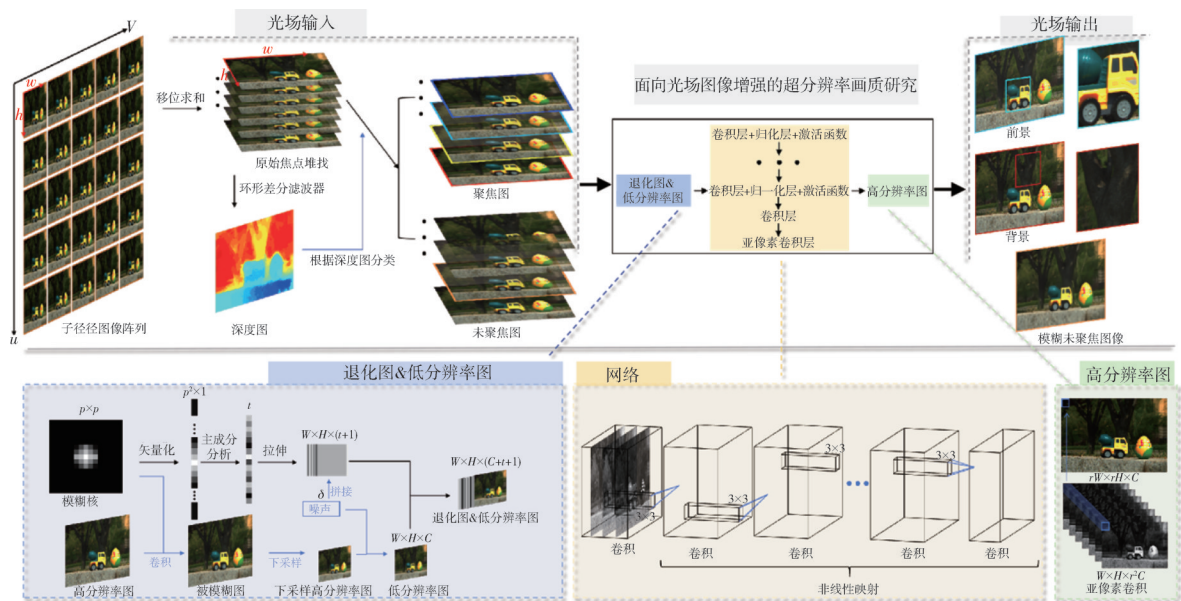


图 2 网络框架

Fig. 2 Network framework

最后,在末层卷积后使用亚像素卷积层,将尺寸为 $W \times H \times r^2 C$ 的多个高分辨率焦点堆栈子图像转换为单个高分辨率焦点堆栈图像,尺寸为 $rW \times rH \times C$ 。对于缩放因子4,网络包含12个卷积层,每层特征图数量设置为128。

2.2 维度拉伸模型

为从低分辨率图像估计高分辨率对应图像,本文求解以下最大后验问题。

$$\hat{x} = \arg \min_x \frac{1}{2\delta^2} \left\| (x \otimes k) \downarrow_r - y \right\|^2 + \lambda \varphi(x), \quad (1)$$

式中: $1/2\delta^2 \left\| (x \otimes k) \downarrow_r - y \right\|^2$ 为数据保真项; $\varphi(x)$

为正则化项(或前一项); λ 为权衡参数,从式(1)可知,估计的解不仅要符合退化过程,而且需要具有高分辨率图像的预期特性; y 为低分辨率图像; k 为模糊核; δ 为噪声水平。由于三者维度不同,直接通过卷积神经网络建模 $\hat{x} =$

$\arg \min_x \frac{1}{2\delta^2} \left\| (x \otimes k) \downarrow_r - y \right\|^2 + \lambda \varphi(x)$ 较为困难。

为此,本文提出如图1所示的维度拉伸策略。假设输入包含尺寸为 $p \times p$ 的模糊核、噪声水平 δ 和尺寸为 $W \times H \times C$ 的低分辨率图像(C 为通道数),处理流程如下:首先,将模糊核向量化为 $p^2 \times 1$ 的向量,并通过主成分分析(Principal Component Analysis, PLA)投影到 t 维线性空间;随后,将降维后的向量与噪声水平 v 拼接,拉伸为尺寸 $W \times H \times (t+1)$ 的退化图 M ,其中所有元素均为 v_i 。在此框架中,低分辨率图像的退化模型可表示为

$$y = (x \otimes k) \downarrow_r + n, \quad (2)$$

式中: $(x \otimes k)$ 为模糊核 k 与潜在高分辨率图像 x 的卷积; $\downarrow_r$ 为缩放因子 r 的下采样操作; n 为标准差为 δ 的加性高斯白噪声。对潜在高分辨率图像进行模糊核卷积得到模糊图像,再对模糊图像进行双三次下采样^[31],最后添加噪声,从而生成成本网络所需低分辨率图像。该过程将退化图与低分辨率图像拼接,使得卷积神经网络能够处理3个输入。

2.3 焦点堆栈生成深度图

在焦点堆栈生成深度图之前,需要通过微透镜阵列解码获取原始光场的全向辐射信息,建立视角间的几何对应关系;在位移加和滤波阶段,依据虚拟成像平面位置调整各子孔径图像的位移量,最终融合生成具有深度连续性的焦点堆栈以模拟不同焦平面的光学聚焦效果。

在光场子孔径图像预处理过程中,对于规则排列的 $N \times N$ 视角阵列输入图像,首先构建包含视角分辨率、空间分辨率及色彩通道的五维光场张量结构 $LF \in \mathbb{R}^{N \times N \times H \times W \times C}$,其中, N 为视角分辨率,对应子孔径图像的网格化空间分布; H 、 W 为子孔径图像的空间分辨率,表征单视角图像细节; C 为颜色通道,兼容多光谱信息。在光场中通过视角相关位移补偿场景视差,使特定深度平面在叠加后转化为参数空间内的连续映射问题。对于 $N \times N$ 规则排列的子孔径图像,第 (u, v) 视角的位移量计算公式为

$$\Delta x(u) = \alpha \left(u - \frac{N+1}{2} \right) s \frac{W}{A_x}, \quad (3)$$

$$\Delta y(v) = \alpha \left(v - \frac{N+1}{2} \right) s \frac{H}{A_y}, \quad (4)$$

式中: α 为基础位移量,由光场相机的几何构型(基线距离 B 与主透镜焦距 f)决定, $\alpha = B/(\tan(\theta/2))$, θ 为视角范围; A_x 、 A_y 为角度-空间轴宽高比补偿系数; $W \times H$ 为空间分辨率;中心视角坐标通常定义为 $(N+1)/2$ 。子孔径图像阵列的补偿逻辑是确保位移在 x 轴和 y 轴方向的比例正确,计算每个子孔径图像在合成焦点堆栈时的位移量,补偿不同焦平面的视差。

基础位移量构建深度与对位移后的坐标 $(x + \Delta x, y + \Delta y)$ 进行重采样。

$$I_{u,v}^{\text{shift}}(x, y) = \text{interp}(I_{u,v}, x + \Delta x, y + \Delta y), \quad (5)$$

式中: interp 为三次立方卷积插值。采用三次立方卷积插值法旨在保持边缘平滑性和连续性,利用周围16个像素拟合连续函数,抑制锯齿伪影。通过插值保持图像连续性,再对所有子孔径图像加权平均,生成不同焦平面的清晰度分布,从原始图像中提取位移后的像素值。 Δx 、 Δy 为由深度信息导出的坐标偏移量,用于模拟不同焦平面下的视差位移(即视差效应)。在光场成像中,物体深度与子孔径图像的视差位移量成反比。深度越大(物体越近),视差位移量 Δx 、 Δy 越小;深度越小(物体越远),位移量越大。

传统插值在图像边缘区域因邻域像素不足(如靠近边界),导致权重和小于1,插值后亮度降低,归一化分母强制权重总和为1,从而确保插值后信号强度在边缘区域与中心区域一致。通过设计更优的插值核函数,减少频域混叠效应,抑制高频混叠,并且归一化操作消除插值过程中因权重不均导致的边缘衰减,保持边缘锐度,该方案在频域传递函数中展现出更优的带限特性。插值后的目标像素值为

$$I_{u,v}^{\text{shift}}(x,y) = \frac{\sum_{u=1}^N \sum_{v=1}^N M(u,v) I_{u,v}^{\text{shift}}(x,y)}{\sum_{u=1}^N \sum_{v=1}^N M(u,v)}. \quad (6)$$

为避免固定核函数的局限性,本文根据局部信号特征动态调整插值权重,改进后的插值核函数与偏移量 u,v 有关,分母对插值结果进行归一化操作,通过分母项平衡不同区域的权重分布确保权重总和为1,保持插值后信号的全局一致性来消除边缘衰减效应。

在焦点堆栈生成阶段,本文设计权重分配机制对高频细节进行强化处理,增强高频细节分离能力,提升深度标签的准确性。生成焦点堆栈目标是为深度估计提供数据基础。首先,定义归一化参数 $\alpha \in [-1, 1]$,表示虚拟焦平面的相对深度位置($\alpha = -1$ 为最远焦平面, $\alpha = +1$ 为最近焦平面),通过遍历 α 参数空间生成焦点堆栈序列 $\{I_{ak}|k=1, \dots, K\}$,以步长 $\Delta\alpha=2/K-1$ 均匀采样,生成 K 层焦点图。本文设计了一种环形差分滤波器(Ring Differential Filter, RDF),其核心由半径递增的同心圆环构成,每个环内集成方向微分滤波器组,该结构通过计算同心环区域内的梯度能量差异,建立像素级清晰度量模型:外环宽尺度滤波捕捉宏观轮廓响应,内环精细微分强化微观细节,中间环层则通过差分运算消除光照变化干扰。进一步地,通过分析焦点堆栈序列中每个像素清晰度随焦距变化的极值特性,构建焦平面位置与物距的动态映射关系:当某像素在多帧焦点堆栈中达到最大清晰度时,其对应的对焦参数可直接转换为深度信息。RDF滤波器结合内圆盘(半径 $r_{\text{in}}=in+gap-1$)、中间间隙(宽度= gap)和外环(半径 $r_{\text{out}}=in+gap+out-1$),数学定义为

$$RDF = \begin{cases} 1, & \sqrt{x^2+y^2} \leq r_{\text{in}}, \\ 0, & r_{\text{in}} < \sqrt{x^2+y^2} \leq r_{\text{in}} + gap, \\ +1, & \text{其他。} \end{cases} \quad (7)$$

归一化后中心点权重调整为 $-\sum RDF$,确保滤波器均值为零。然后进行聚焦度量计算,对于焦点堆栈中第 k 帧图像,其聚焦能量 E_k 为

$$E_k(x,y) = \sum_{c \in \{R, G, B\}} (I_k^{(c)} \otimes RDF)^2(x,y). \quad (8)$$

能量值 E_k 越大,表明像素 (x,y) 在第 k 帧越清晰。初始深度标签 L_{init} 通过最大化焦点堆栈帧的聚焦点能量确定。

$$L_{\text{init}}(x,y) = \underset{1 \leq k \leq N}{\operatorname{argmax}} E_k(x,y). \quad (9)$$

初始全聚焦图像通过选择性整合最优聚焦像

素生成,

$$AIF_{\text{init}}(x,y) = L_{\text{init}}(x,y). \quad (10)$$

为抑制噪声并保留边缘,应用引导滤波。滤波后能量更新为

$$E_k = aE_k + b. \quad (11)$$

为抑制噪声并保持边缘,采用引导滤波对初始深度标签进行优化,以全聚焦图像作为引导,滤波后的能量为 $E_k = aE_k + a$,其中,线性系数 a,b 通过最小化局部窗口 ω 内的均方误差确定,

$$\min_{a,b} \sum_{(i,j) \in \omega} (aE_k(i,j) + b - L_{\text{init}}(i,j))^2 + \epsilon a^2. \quad (12)$$

优化后重新选择深度标签 L_{filtered} ,并生成平滑后的全聚焦图像 AIF_{filtered} 。进一步结合空间邻域与颜色一致性,定义权重函数 $\omega(p,q)$ 为

$$\omega(p,q) = \exp\left(-\frac{\|AIF(p) - AIF(q)\|^2}{\delta^2}\right) \exp\left(-\frac{\|p - q\|^2}{r^2}\right). \quad (13)$$

对每个像素 p ,在邻域 Ω_p 内计算加权中值,得到最终深度标签 L_{final} 为

$$L_{\text{final}}(p) = \operatorname{median}\{L_{\text{filtered}}(q) | q \in \Omega_p, \omega(p,q) > r\}. \quad (14)$$

然后进行深度图归一化并得到最终输出,将深度标签 L_{final} 线性映射至 $[0, 255]$ 范围,生成灰度深度图为

$$DMAP(x,y) = \frac{L_{\text{final}}(x,y) - \min(L_{\text{final}})}{\max(L_{\text{final}}) - \min(x,y)} \times 255. \quad (15)$$

最后,深度图被线性缩放为8位灰度图。该方法通过多级优化框架平衡噪声抑制与边缘保留,有效从焦栈中恢复场景深度信息。

3 训练

对于模糊核,本文采用各向同性高斯核模型(固定核宽),其在实际应用中已被证明可行。缩放因子为4时,核宽范围设为 $[0.2, 4]$,步长0.1采样。核尺寸固定为 15×15 。为进一步扩展退化空间,本文还考虑各向异性高斯核模型,其参数由特征向量的旋转角和特征值缩放决定。旋转角范围设为 $[0, \pi]$,缩放因子为4时,特征值缩放设为10。当模糊核定义明确或已学习时,本文均匀采样多种核并通过PCA投影矩阵聚合。通过保留约99.8%的能量,核被投影到15维空间(即 $t=15$)。噪声水平范围设为 $[0,$

75]。训练时, 每轮随机选择模糊核和噪声水平合成 LR 图像, 并裁剪 $N=128 \times 3\,000$ 个 LR/HR 图像块对(含退化图)。批量大小为 128, 初始学习率为 10^{-3} , 当训练误差停止下降时降至 10^{-4} 。若误差连续 5 轮未变化, 将批量归一化层参数合并至相邻卷积滤波器, 随后以 10^{-5} 的小学习率微调 100 轮。

4 结果与分析

本文在公开和自建光场数据集上进行了实验。采用 3 个公开数据集为 EPFL^[32]、INRIA^[33] 和 STFgantry^[3]。自建数据集使用 Lytro Illum 光场相机(1 英寸传感器, F/2.0 光圈, 8 倍光学变焦, 焦距范围 30~240 mm)。训练、验证和测试集分别包含 250、50 和 200 张光场图像, 所有数据预处理为 9×9 角分辨率并裁剪为更小的角分辨率。将相同裁剪尺寸的子孔径图像转换为焦点堆栈进行训练。图像通过双三次插值下采样 4 倍, 并使用本方法进行超分辨率处理。由于缺乏高分辨率图像作为地面真实值计算峰值信噪比, 本文采用图像信息熵和高频能量评估焦点堆栈的超分辨率性能。所有算法在配备 2.40 GHz

Intel Core i7-13700 CPU 和 16GB 内存的计算机上通过 Matlab 实现。

4.1 定性评估

本文所提算法与 Liu 等^[26]提出的利用视角内外交互的网络处理、Wang 等^[27]提出的视差估计进行比较。从主观角度评价图像质量, 对于真实光场图像, 本文采用 Lytro Illum 相机(1 英寸传感器, F/2.0 光圈, 8 倍变焦)拍摄图像并建立自开发光场数据集, 训练集、验证集和测试集分别包含 250、50 和 200 张光场图像, 子孔径图像经过预处理, 形成 9×9 角分辨率并裁剪为更小尺寸后转换为焦点堆栈图像。实际拍摄过程中, 为获得较为清晰可靠的光场图像, 采用三脚架作为辅助工具。原始数据通过 Matlab 的 LF Toolbox^[30]处理得到子孔径图像集。训练集包含多样化场景。如图 3~图 6 所示, 本方法利用不同焦平面图像的清晰度信息估计像素深度, 并从焦点堆栈图像中提取前景与背景。最终将同一焦平面生成的前景与背景图像与其他方法对比。结果显示, 本方法的聚焦图像细节更精细, 主体更突出。

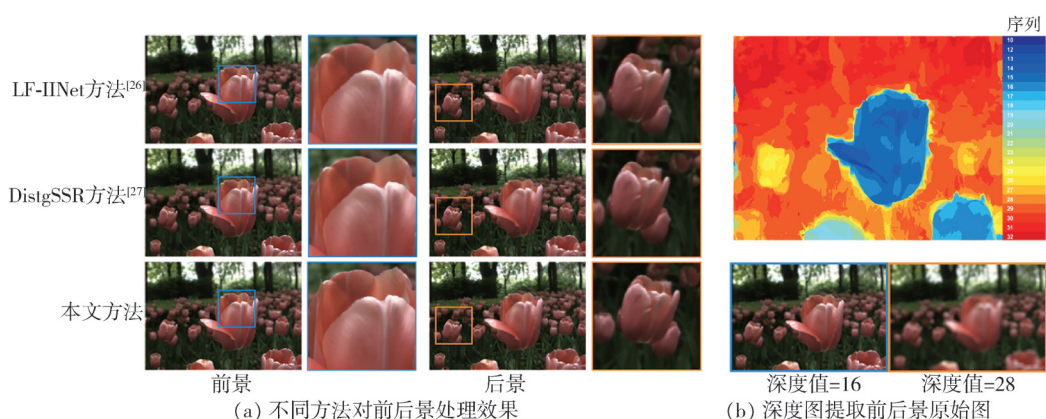


图 3 不同方法对郁金香前后景超分辨率对比图
Fig. 3 Comparison chart of SR image of tulip foreground and background using different methods

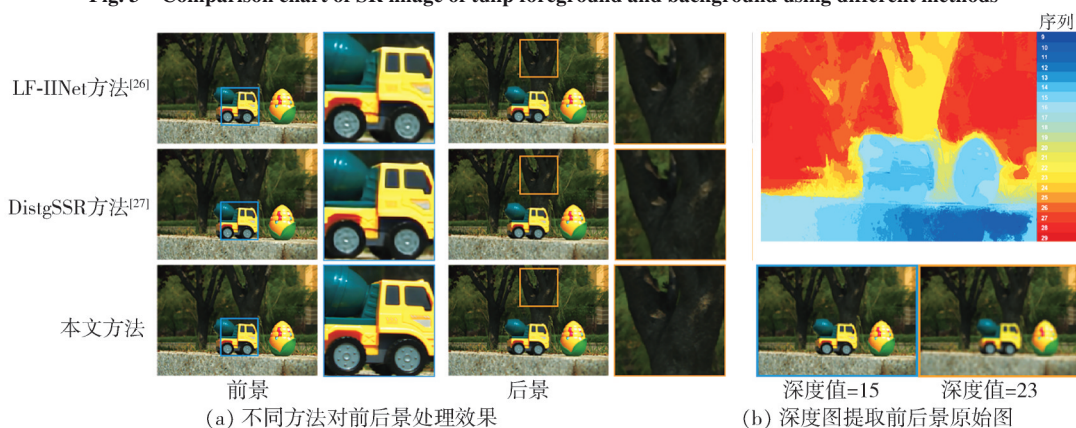


图 4 不同方法对玩具前后景超分辨率对比图
Fig. 4 Comparison chart of SR image of toys foreground and background using different methods

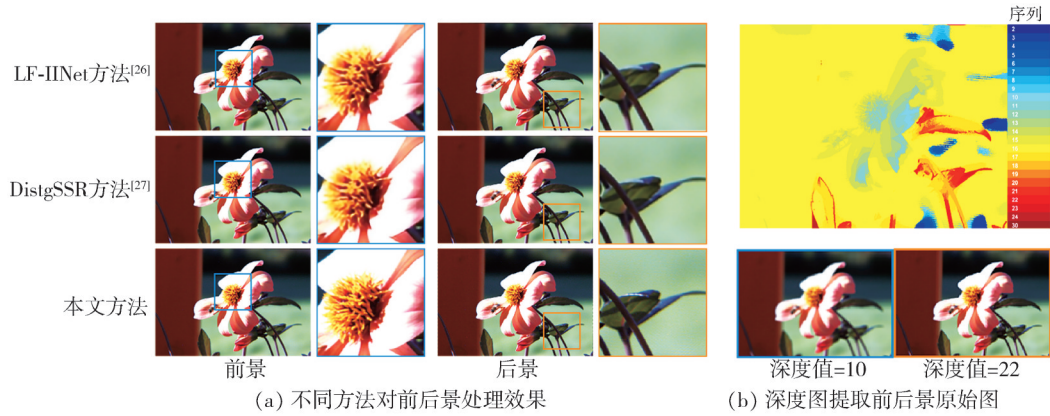


图 5 不同方法对花前后景超分辨率对比图

Fig. 5 Comparison chart of SR image of flower foreground and background using different methods

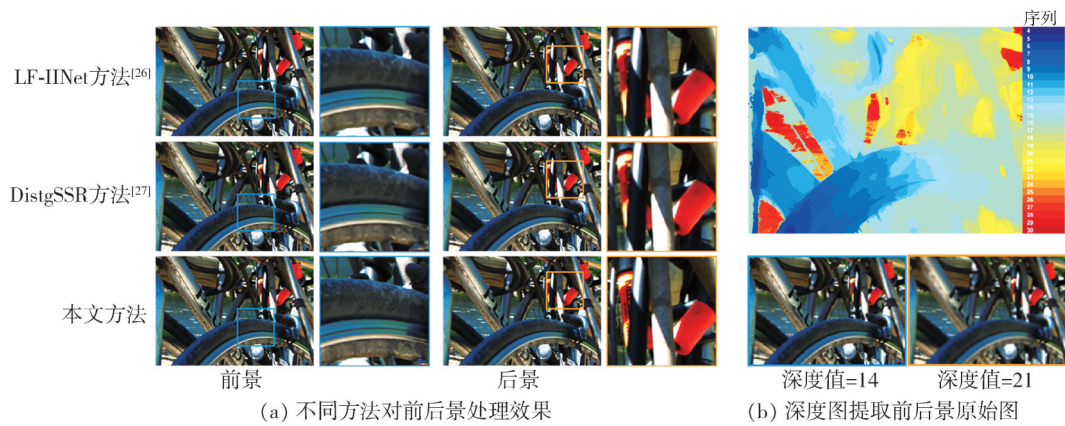


图 6 不同方法对车轮前后景超分辨率对比图

Fig. 6 Comparison chart of SR image of wheel foreground and background using different methods

4.2 定量评估

本文对比了两种基于深度学习的方法: LF-IINet^[26]和 Distg-SSR^[27]。表 1 和表 2 展示了不同

方法在各验证和测试场景中的定量结果。在 4 个数据集上, 本方法在信息熵评估中表现最优。在高频能量评估中, 本文方法在全部场景的前景和背景中均位列第一。

表 1 图像信息熵定量评价

Tab. 1 Quantitative evaluation of information entropy

方法		EPFL 中图像信息熵		INRIA 中图像信息熵		STFgantry 中图像信息熵		自建数据集中图像信息熵	
		Paved Road	Sphynx	Building	Sculpture	Lego Knights	Tarot CardsS	Flower02	Podium5D
LF-IINet	前景	7.367 1	7.173 2	7.274 8	7.548 0	7.201 0	7.339 6	6.980 3	6.625 1
	后景	7.367 8	7.181 4	7.275 9	7.547 5	7.236 2	7.314 1	6.991 3	6.641 8
Distg-SSR	前景	7.358 4	7.169 2	7.269 7	7.550 9	7.199 6	7.339 1	6.979 8	6.639 0
	后景	7.359 0	7.177 7	7.268 2	7.550 4	7.236 2	7.315 1	6.990 2	6.627 8
本文方法	前景	7.410 5	7.191 7	7.285 3	7.555 9	7.213 2	7.345 9	6.982 9	6.641 8
	后景	7.431 7	7.205 3	7.298 1	7.555 4	7.238 2	7.316 1	6.991 0	6.651 7

表 2 图像高频能量定量评价

Tab. 2 Quantitative evaluation of high-frequency energy

方法		EPFL 中高频能量		INRIA 中高频能量		STFgantry 中高频能量		自建数据集中高频能量	
		Paved Road	Sphynx	Building	Sculpture	Lego Knights	Tarot CardsS	Flower02	Podium5D
LF-IINet	前景	2.512 9	2.658 6	3.783 9	1.612 1	0.777 3	2.476 9	0.644 3	0.747 5
	后景	2.556 6	2.519 2	4.230 2	1.604 8	0.799 7	1.549 0	0.483 5	0.758 1
Distg-SSR	前景	2.476 5	2.650 1	3.815 6	1.619 4	0.781 8	2.511 4	0.652 5	0.766 9
	后景	2.536 3	2.518 8	4.295 3	1.610 3	0.795 9	1.562 5	0.484 1	0.780 5
本文方法	前景	3.146 5	2.891 0	4.004 6	1.926 5	0.815 6	2.806 0	0.672 0	0.793 9
	后景	3.530 3	2.735 8	4.669 6	1.950 0	0.814 1	1.585 0	0.487 9	0.827 5

结果表明所提方法的图像信息熵较LF-IINet方法平均提升前景0.21%，平均提升后景0.22%，较Distg-SSR方法平均提升前景0.21%，平均提升后景0.28%；在图像高频能量方面，所提方法较LF-IINet方法平均提升前景10.78%，平均提升后景10.56%，较Distg-SSR方法平均提升前景11.33%，平均提升后景11.02%。在对比实验分析中，本文方法展现出显著技术优势，可以充分利用焦点堆栈的深度信息，处理场景深度突变区域的高频细节，突出显示深度变化和细节信息，使得主体部分的细节和边缘更加清晰。

5 结 论

本文提出将光场原始子孔径图像集转换为焦点堆栈图像，利用不同焦平面图像的清晰度信息估计像素深度，并根据深度图将焦栈分为聚焦与未聚焦图像。对聚焦图像的前景和背景进行选择超分辨率处理，未聚焦图像可添加不同模糊核的模糊或保持原状。在公开和自建光场数据集上进行了实验，结果表明所提方法的图像信息熵较LF-IINet方法平均提升前景0.21%，平均提升后景0.22%，较Distg-SSR方法平均提升前景0.21%，平均提升后景0.28%；在图像高频能量方面，所提方法较LF-IINet方法平均提升前景10.78%，平均提升后景10.56%，较Distg-SSR方法平均提升前景11.33%，平均提升后景11.02%。通过选择性超分辨率的网络生成高分辨率焦栈聚焦图像和模糊未聚焦图像，有效提升了主体显示效果。

利益冲突声明/Conflict of Interests

所有作者声明不存在利益冲突。

All authors declare no relevant conflict of interests.

参考文献

- [1] 刘严羊硕, 潘晋孝, 潘毅华, 等. 三维场景的光场成像仿真[J]. 测试技术学报, 2018, 32(2): 119-124.
LIU Yanyangshuo, PAN Jinxiao, PAN Yihua, et al. Light field photography simulation of three-dimensional scene[J]. Journal of Test and Measurement Technology, 2018, 32(2): 119-124. (in Chinese)
- [2] ZHU H, GUO M, LI H, 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.
- [3] NG R, LEVOY M, BRÉDIF M, et al. Light field photography with a hand-held plenoptic camera[D]. Palo Alto: Stanford University, 2005.
- [4] VENKATARAMAN K, LELESCU D, DUPARRÉ J, et al. Picam: an ultra-thin high performance monolithic camera array[J]. ACM Transactions on Graphics (TOG), 2013, 32(6): 1-13.
- [5] WANG Y, YANG J, GUO Y, et al. Selective light field refocusing for camera arrays using bokeh rendering and superresolution[J]. IEEE Signal Processing Letters, 2019, 26(1): 204-208.
- [6] 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.
- [7] WANG T C, EFROS A A, RAMAMOORTHY R. Depth estimation with occlusion modeling using light-field cameras[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(11): 2170-2181.
- [8] WILLIEM, PARK I K, LEE K M. Robust light field depth estimation using occlusion-noise aware data costs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(10): 2484-2497.
- [9] JEON H G, PARK J, CHOE G, et al. Depth from a light field image with learning-based matching costs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(2): 297-310.
- [10] ZHOU M, DING Y, JI Y, et al. Shape and reflectance reconstruction using concentric multi-spectral light field[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(7): 1594-1605.
- [11] WANG T C, CHANDRAKER M, EFROS A A, et al. SVBRDF-invariant shape and reflectance estimation from a light-field camera[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(3): 740-754.
- [12] TAO M W, SU J C, WANG T C, et al. Depth estimation and specular removal for glossy surfaces using point and line consistency with light-field cameras[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(6): 1155-1169.
- [13] ALPEROVICH A, JOHANNSEN O, STRECKE M, et al. Light field intrinsics with a deep encoder-decoder network[C]//2018 IEEE/CVF Conference

- on Computer Vision and Pattern Recognition, 2018: 9145-9154.
- [14] KIM C, ZIMMER H, PRITCH Y, et al. Scene reconstruction from high spatio-angular resolution light fields[J]. ACM Transactions on Graphics, 2013, 32(4): 1-12.
- [15] LI N, YE J, JI Y, et al. Saliency detection on light field[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(8): 1605-1616.
- [16] SRINIVASAN P P, NG R, RAMAMOORTHY R. Light field blind motion deblurring[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 2354-2362.
- [17] SCHAMBACH M, HEIZMANN M. A multispectral light field dataset and framework for light field deep learning[J]. IEEE Access, 2020, 8: 193492-193502.
- [18] 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.
- [19] 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), 2015: 57-65.
- [20] YOON Y, JEON H G, YOO D, et al. Light-field image super-resolution using convolutional neural network[J]. IEEE Signal Processing Letters, 2017, 24(6): 848-852.
- [21] WU G, ZHAO M, WANG L, et al. Light field reconstruction using deep convolutional network on EPI [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 1638-1646.
- [22] WU G, LIU Y, DAI Q, et al. Learning sheared EPI structure for light field reconstruction[J]. IEEE Transactions on Image Processing, 2019, 28(7): 3261-3273.
- [23] WU G, WANG Y, LIU Y, et al. Spatial-angular attention network for light field reconstruction [J]. IEEE Transactions on Image Processing, 2021, 30: 8999-9013.
- [24] 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, 2018: 4748-4757.
- [25] ZHANG S, LIN Y, SHENG H. Residual networks for light field image super-resolution[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 11038-11047.
- [26] LIU G, YUE H, WU J, et al. Intra-inter view interaction network for light field image super-resolution [J]. IEEE Transactions on Multimedia, 2021, 25: 256-266.
- [27] WANG Y, WANG L, WU G, et al. Disentangling light fields for super-resolution and disparity estimation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(1): 425-443.
- [28] ZHANG K, LIANG J, VAN GOOL L, et al. Designing a practical degradation model for deep blind image super-resolution[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021: 4771-4780.
- [29] LIPPMANN G. Épreuves réversibles donnant la sensation du relief[J]. Journal de Physique Théorique et Appliquée, 1908, 7(1): 821-825.
- [30] LIANG C K, LIN T H, WONG B Y, et al. Programmable aperture photography [J]. ACM Transactions on Graphics, 2008, 27(3): 1-10.
- [31] CHUNG M, JUNG M, KIM Y. Enhancing remote sensing image super-resolution guided by bicubic-downsampled low-resolution image[J]. Remote Sensing, 2023, 15(13): 3309.
- [32] RERABEK M, EBRAHIMI T. New light field image Dataset [C]//2016 8th International Conference on Quality of Multimedia Experience (QoMEX), 2016: 1-8.
- [33] MAGGIORI E, TARABALKA Y, CHARPIAT G, et al. Can semantic labeling methods generalize to any city? The inria aerial image labeling benchmark[C]//2017 IEEE International geoscience and remote sensing symposium (IGARSS), 2017: 3226-3229.