

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

角度差异强化的光场图像超分网络

吕天琪, 武迎春, 赵贤凌

引用本文:

吕天琪, 武迎春, 赵贤凌. 角度差异强化的光场图像超分网络[J]. 光电工程, 2023, 50(2): 220185.

Lv T Q, Wu Y C, Zhao X L. Light field image super-resolution network based on angular difference enhancement[J]. *Opto-Electron Eng*, 2023, 50(2): 220185.

<https://doi.org/10.12086/oe.2023.220185>

收稿日期: 2022-07-28; 修改日期: 2022-10-17; 录用日期: 2022-10-21

相关论文

基于边框加权角相关的光场深度估计算法

马帅, 王宁, 朱里程, 王帅, 杨平, 许冰

光电工程 2021, 48(12): 210405 doi: 10.12086/oe.2021.210405

基于光场内联遮挡处理的噪声场景深度获取

吴迪, 张旭东, 范之国, 孙锐

光电工程 2021, 48(7): 200422 doi: 10.12086/oe.2021.200422

融合多尺度特征的光场图像超分辨率方法

赵圆圆, 施圣贤

光电工程 2020, 47(12): 200007 doi: 10.12086/oe.2020.200007

更多相关论文见光电期刊集群网站 

 | 光电工程
Opto-Electronic Engineering

<http://cn.ojournal.org/oe>



 OE_Journal

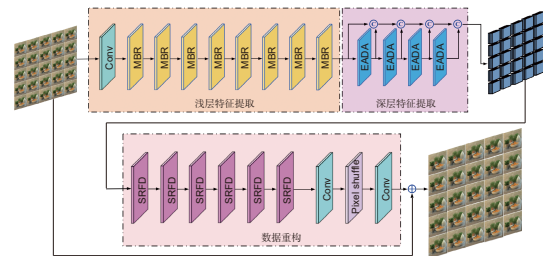


Website

DOI: 10.12086/oe.2023.220185

角度差异强化的光场图像超分网络

吕天琪, 武迎春*, 赵贤凌

太原科技大学电子信息工程学院,
山西 太原 030024

摘要: 由于采用了更为先进的成像技术, 光场相机可以同步获取场景的空间信息与角度信息。该技术以牺牲空间分辨率为代价, 实现了更高维度的场景表示。为了提高光场相机拍摄场景的空间分辨率, 本文搭建了角度差异强化的光场超分辨率重构网络。该网络先采用 8 个多分支残差块实现浅层特征提取, 再采用 4 个强化的角度可变形对准模块实现深层特征提取, 最后采用 6 个简化的残差特征蒸馏模块和像素洗牌模块完成数据重构。所提网络在利用光场角度差异完成空间信息超分时, 更加强调视图自身特征的深入挖掘, 以获得更加丰富的视图间差异特征。在 5 组公开的光场数据集上对本文所提网络的性能进行了验证, 本文算法获得了 PSNR、SSIM 值更高的高分辨率光场子孔径图像。

关键词: 光场相机; 超分辨率重构; 残差; 可变形卷积; 残差特征蒸馏

中图分类号: TP391.4

文献标志码: A

吕天琪, 武迎春, 赵贤凌. 角度差异强化的光场图像超分网络 [J]. 光电工程, 2023, 50(2): 220185

Lv T Q, Wu Y C, Zhao X L. Light field image super-resolution network based on angular difference enhancement[J]. *Opto-Electron Eng*, 2023, 50(2): 220185

Light field image super-resolution network based on angular difference enhancement

Lv Tianqi, Wu Yingchun*, Zhao Xianling

School of Electronic and Information Engineering, Taiyuan University of Science and Technology, Taiyuan, Shanxi 030024, China

Abstract: Based on the advanced imaging technology, light field camera can obtain the spatial information and the angular information of the scene synchronously. It achieves higher dimensional scene representation by sacrificing the spatial resolution. In order to improve the spatial resolution of the light field image, a light field super-resolution reconstruction network based on angle difference enhancement is built in this paper. In the proposed network, eight multi-branch residual blocks are used to extract shallow features. Then, four enhanced angular deformable alignment modules are used to extract deep features. Finally six simplified residual feature distillation modules and pixel shuffle modules are used to complete data reconstruction. The proposed network takes advantage of the angle difference of the light field to complete the spatial information super-resolution. In order to obtain more features difference between different views, the own feature of the single view is emphasized during the feature

收稿日期: 2022-07-28; 修回日期: 2022-10-17; 录用日期: 2022-10-21

基金项目: 国家自然科学基金资助项目 (61601318); 山西省基础研究计划资助项目 (202103021224278); 山西省回国留学人员科研资助项目 (2020-128)

*通信作者: 武迎春, yingchunwu3030@foxmail.com。

版权所有©2023 中国科学院光电技术研究所

extraction. The performance of the proposed network is verified on five public light field data sets. The proposed algorithm obtains high-resolution light field sub-aperture images with higher PSNR and SSIM.

Keywords: light field camera; super-resolution reconstruction; residual; deformable convolution; residual feature distillation

1 引言

传统相机记录 3D 空间的 2D 投影, 照片的强度反映了空间光线的能量积分。为了更加充分地记录空间信息, 常通过变换相机位置及调整镜头焦距, 实现多视角图像及多聚焦图像的获取^[1]。与传统相机相比, 光场相机的成像效率更高, 基于单次曝光捕获的光场原图, 通过图像处理即可计算同一场景的多张重聚焦图像及多视角图像^[2]。光场相机成像效率的提升得益于相机成像模型的优化^[3], 核心技术在于通过分光技术实现 4D 光场的获取, 即空间光线位置信息和角度信息的同步获取, 但角度信息的获取是以牺牲图像空间分辨率为代价的。为了提高光场图像的空间分辨率, 研究有效的光场超分辨率重构技术具有重要意义^[4-11]。近年来, 各种先进的深度学习网络框架被应用于图像超分辨率重构, 同时在光场超分辨率重构中也展示了其优越的性能。悉尼大学 Yeung 等人提出空间、角度信息可分离的光场超分 (light field super-resolution using deep efficient spatial-angular separable convolution, LFSSR) 网络, 该模型呈沙漏形状, 设计了空间-角度可分离卷积来提取空间与角度的联合特征, 具有运算速度快的优点^[7]。北京交通大学张硕等人提出残差光场超分 (residual networks for light field image super-resolution, ResLF) 网络, 该方法隐式地挖掘多视点图像间的内在对应关系, 通过学习子孔径图像阵列中水平、竖直和对角方向上的视差信息, 实现了重建图像的高频信息补充^[8]。但 ResLF 算法由于没有用到光场中的全部视角图像, 从而导致对遮挡边缘部分的重建结果有些失真。香港城市大学 Jin 等人提出结合几何嵌入与结构一致性的光场超分网络, 首先利用多对一模块学习子孔径图像间的互补信息, 得到初级超分图像, 再采用正则化模块强化视差结构, 得到空间结构一致性增强的高分辨率图像^[9]。网络的优点在于充分利用视图之间的互补信息和保留光场视差结构。以宏像素图像阵列为输入, 国防科技大学王应谦等人提出空间、角度交互光场超分 (spatial-angular interaction for light field image super-resolution, LF-InterNet) 网络,

提出空间特征提取器与角度特征提取器分别提取光场图像的空间特征与角度特征, 将其反复交互融合得到高分辨率光场图像^[4]。并且空间角度交互机制是一个通用框架, 可以高效地处理光场数据。之后, 他们又提出了可变形卷积光场图像超分 (light field image super-resolution using deformable convolution, LF-DFnet) 网络, 设计了角度可变形对准模块来学习侧视图与中心视图的偏移量, 并将该偏移量编码到每个视图的特征中提高网络超分性能^[5]。可变形卷积光场图像超分网络虽然考虑到了所有视图之间的视差信息, 但是在深层特征提取部分却忽略了对自身特征深入学习。国防科技大学梁政宇等人提出角度分辨率灵活变化的光场超分 (angular-flexible network for light field image super-resolution, LF-AFnet) 网络, 将解耦和融合模块以残差的形式连接, 实现了任意角度分辨率光场图像的超分辨率重构^[10]。其最大的优势在于该网络可以处理具有任意角度分辨率的不同类型设备捕获的光场数据。国防科技大学莫宇等人提出稠密双注意力网络 (dense dual-attention network, DDAN), 设计了视图注意力模块和通道注意力模块, 实现了跨视点和跨通道有用信息的提取, 提高了超分过程中子孔径图像信息的利用率^[6]。其优点在于能够以更少的网络参数量达到更好的超分辨率效果。上海交通大学赵圆圆等人提出融合多尺度特征的光场图像超分辨率网络^[11], 该算法在空洞空间金字塔池化 (atrous spatial pyramid pooling, Asp) 模块结构基础上加入了残差式的设计, 组成了残差空洞空间金字塔池化 (residual atrous spatial pyramid pooling, ResAsp) 的网络子结构, 该算法没有充分利用视图的差异特征。

光场超分网络以 4D 光场重排后的 2D 图像阵列为输入, 大多通过学习光场子孔径图像间的空间特征和角度特征来进行超分得到高分辨率图像。性能优越的 LF-DFnet 网络在学习过程中引入了不同视图间差异特征的学习, 却忽略了视图自身特征的学习, 而在传统单幅图像超分辨率重建过程中, 通过图像自身特征学习也可得到质量较高的高分辨率图像。基于此, 本文搭建了自身特征强化的光场超分辨率重构网络,

采用多分支残差块实现浅层特征提取, 在深层特征提取时采用优化的角度可变形对准模块来强化中心视图特征和侧视图特征的学习, 最后采用简化的残差特征蒸馏模块实现子孔径图像重构。

2 光场超分辨率重构网络搭建

基于光场双平面表示模型如图 1(a) 所示, 空间任一光线可由穿过(u, v)平面和(s, t)平面的两个点确定, 由光场相机捕获的 4D 光场由 $L_F(s, t, u, v)$ 表示, 其中 $s \in [1, S], t \in [1, T], u \in [1, U], v \in [1, V], U \times V$ 代表 4D 光场的角度分辨率, $S \times T$ 表示 4D 光场的空间分辨率, 一般情况下, $U = V = A$ 。为了实现 4D 光场可视化, 常将 4D 光场重排为子孔径图像阵列或宏像素阵列, 如图 1(b) 和 1(c) 所示, 一般光场超分辨率重建网络以这两种数据格式为输入。

考虑到光场相机角度信息的获取是以牺牲图像空间分辨率为代价, 本文搭建的光场超分辨率重建网络

主要用于光场图像空间分辨率超分, 整体网络结构图如图 2 所示。网络的输入为 4D 光场重排后的子孔径图像阵列, 该阵列由 1 幅中心视图和 $A^2 - 1$ 幅侧视图组成, 网络包括浅层特征提取、深层特征提取、数据重构三部分。其中浅层特征提取采用 8 个多分支残差 (multi-branch residual, MBR) 块实现, 深层特征提取采用 4 个优化的角度可变形对准 (enhanced angular deformable alignment, EADA) 模块实现, 数据重构中特征融合采用 6 个简化的残差特征蒸馏 (simplified residual feature distillation, SRFD) 模块实现, 特征上采样通过像素洗牌 (pixel shuffle)^[4-7] 完成。

2.1 基于 MBR 块的浅层特征提取

本文搭建的光场超分辨率重建网络在进行浅层特征提取时, 采用 1×1 卷积和 8 个 MBR 块生成浅层特征, 所设计的 MBR 块包括 3 个支路, 其数据处理流程图如图 3 所示, 支路 1 包含 2 个卷积核大小为 3×3

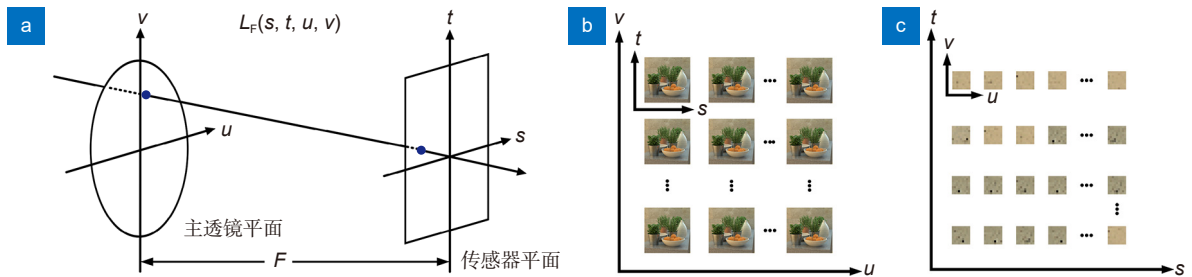


图 1 4D 光场获取与重排。(a) 光场双平面表示模型; (b) 子孔径图像阵列; (c) 宏像素阵列
Fig. 1 4D light field acquisition and rearrangement. (a) Biplanar representation model of light field; (b) Subaperture image array; (c) Macropixel array

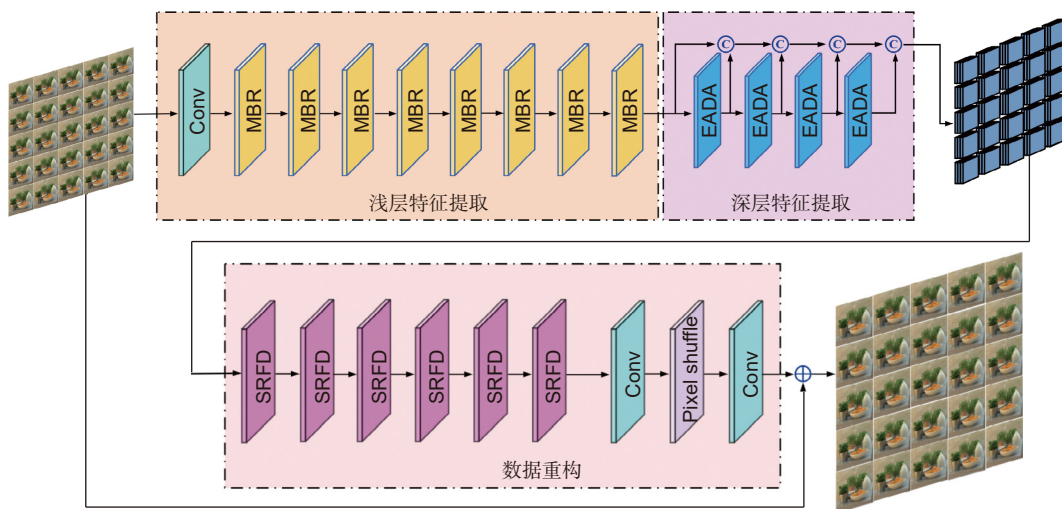


图 2 整体网络结构图
Fig. 2 Overall network structure diagram

的空洞卷积和 2 个卷积核大小为 1×1 普通卷积, 空洞卷积的膨胀因子为 12, 支路 2 和支路 3 通过跳跃连接直接将残差块的输入向后传递, 并且使用跳跃连接的方法还可以将初级特征与高级特征融合, 很好地解决了网络退化问题。使用空洞卷积可以提取丰富的层次信息并扩大了特征感受野。

2.2 基于 EADA 模块的深层特征提取

4D 光场数据较传统 2D 图像, 具有丰富的角度信息, 故光场超分辨率重构网络在深层特征提取时, 有效利用角度信息的相关性可以提高网络超分性能。

文献 [5] 提出的 LF-DFnet 网络, 在深层特征提取部分用到了角度可变形对准 (angular deformable alignment, ADA) 模块, 其结构如图 4(a) 所示, 该模块利用可变形卷积块 (deformable convolution block, DCB) 来进行特征收集和特征发散。特征收集部分如图 4(b) 所示, 数据处理流程包括:

1) 将侧视图特征 1 与中心视图特征对齐, 学习对齐后的侧视图特征 1 与中心视图特征来得到偏移量特征 (如图 4(c) 所示), 将侧视图特征 1 与偏移量特征送入 DCB 进行特征收集, 得到侧视图到中心视图的对

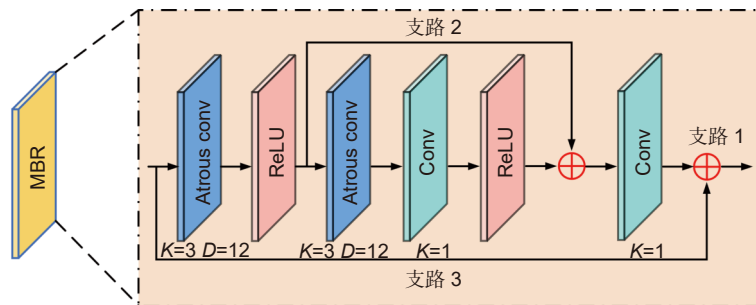


图 3 多分支残差块

Fig. 3 Multi-branch residual block

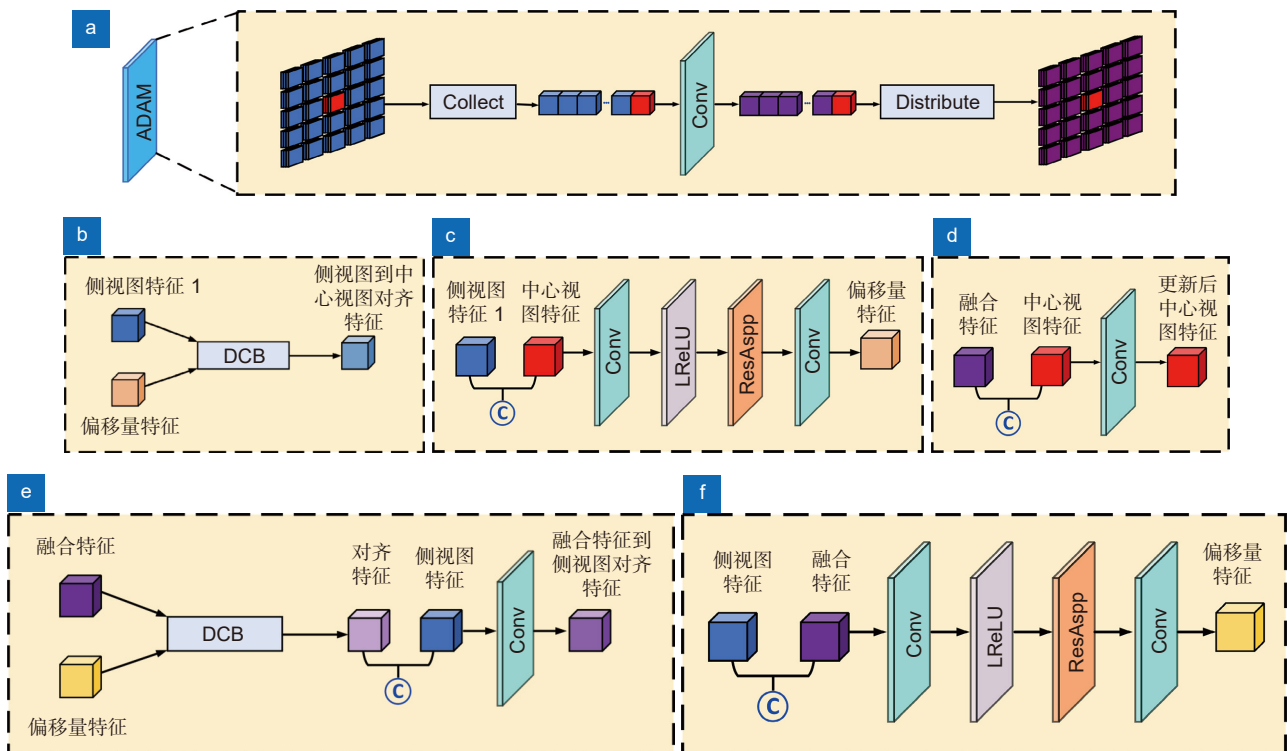


图 4 ADA 模块细节。(a) ADA 模块数据处理流程; (b) 特征收集 (collect); (c) 特征收集集中偏移量的获取; (d) 中心视图的更新; (e) 特征发散 (distribute); (f) 特征发散中偏移量的获取

Fig. 4 ADA module details. (a) ADA module data processing process; (b) Feature collection; (c) Offset acquisition in feature collection; (d) Update the central view; (e) Feature distribution; (f) Offset acquisition in feature distribution

齐特征;

2) 更新中心视图特征 (具体操作如图 4(d) 所示), 将侧视图 2 的特征与中心视图特征对齐, 之后数据处理与过程 1) 相同;

3) 循环收集所有侧视图的偏移量特征;

4) 所有侧视图特征和更新后的中心视图特征通过 1×1 卷积进行角度信息合并, 得到融合特征。

特征发散部分如图 4(e) 所示, 先将融合特征与侧视图特征进行通道拼接, 经卷积、激活、残差空洞空间金字塔池化 (residual atrous spatial pyramid pooling, ResAspp) 等操作学习侧视图特征与融合特征的偏移量 (如图 4(f) 所示), 然后将偏移量特征与融合特征一起送入 DCB 进行特征发散。通过这种方式, 角度信息可以合并到每个侧视图中, 使得所有视角的超分性能得到均匀改善。

为了充分利用 4D 光场多视角图像间角度信息的相关性, 本文在深层特征提取部分也采用可变形卷积来进行特征收集和特征发散。考虑到 ADA 模块在深层特征提取中没有充分对视图自身特征进行挖掘, 本文对 ADA 模块进行了改进, 提出了强化侧视图自身特征学习的 EADA 模块, 以进一步提高网络性能。EADA 模块中特征收集部分的结构如图 5(a) 所示, 与 ADA 模块中的特征收集相比, EADA 模块在利用 DCB 实现特征收集和特征发散之前, 先对每个侧视图特征进行强化学习。具体来说, 在特征收集部分, 先用 1×1 卷积操作提取侧视图特征, 然后再用两支路来进一步特征收集, 支路 1 增加了 2 个卷积核大小为 3×3 的空洞卷积, 空洞卷积的膨胀因子分别为 2 与 4; 支路 2 增加了 1 个卷积核大小为 1×1 普通卷积和 1 个卷积核大小为 3×3 的空洞卷积, 空洞卷积的膨胀因子为 2。然后将强化后的侧视图特征 1 与中心视图特征对齐来得到偏移量特征 1 (过程与图 4(c) 相同), 偏移量特征 1 与强化后的侧视图特征 1 一起馈送入 DCB 模块得到对齐特征 1, 同理得到对齐特征 2。对齐特征 1 和对齐特征 2 可以用式 (1) 和式 (2) 来表示:

$$F_k^{s \rightarrow c} = H_{DCB}([H_{DFE1}(F_{k-1}^s), \Delta P_k^{s \rightarrow c}], \quad (1)$$

$$F_k^{s \rightarrow c} = H_{DCB}([H_{DFE2}(F_{k-1}^s), \Delta P_k^{s \rightarrow c}], \quad (2)$$

其中: $F_k^{s \rightarrow c}$ 表示支路 1 第 $k \in [1 \sim A^2 - 1]$ 幅侧视图 (side-view) 对中心视图 (center-view) 的对齐特征 1, $F_k^{s \rightarrow c}$ 表示支路 2 第 $k \in [1 \sim A^2 - 1]$ 幅侧视图 (side-view) 对中心视图 (center-view) 的对齐特征 2, H_{DCB} 表示可变形卷积块, H_{DFE1} 表示支路 1 的普通卷积和两个空洞卷积

操作, H_{DFE2} 表示支路 2 的普通卷积和空洞卷积操作。 F_{k-1}^s 代表支路 1 强化后的侧视图特征 1, F_{k-1}^s 代表支路 2 强化后的侧视图特征 2。 $\Delta P_k^{s \rightarrow c}$ 和 $\Delta P_k^{s \rightarrow c}$ 分别代表支路 1 和支路 2 的偏移量特征。最后将两个支路得到的对齐特征进行通道拼接后, 经过卷积核大小为 1×1 普通卷积得到对齐特征 3。在特征发散部分, 本文在 DCB 之前添加了 1 个卷积核大小为 1×1 普通卷积和 2 个卷积核大小为 3×3 的空洞卷积, 来对侧视图特征强化学习, 空洞卷积的膨胀因子分别为 2 与 4。侧视图的最终发散特征可以根据式 (3) 来表示:

$$F_k^s = H_{squeeze}^k([H_{DCB}(F_{fuse,k}^s, \Delta P_k^s), F_{k-1}^s]), \quad (3)$$

其中: F_k^s 表示特征发散后的侧视图特征, $H_{squeeze}^k$ 代表通道降维, $F_{fuse,k}^s$ 表示融合特征, ΔP_k^s 代表偏移量特征, F_{k-1}^s 代表最原始的侧视图特征。EADA 模块充分挖掘了各个侧视图的特征信息, 可以很好地将角度信息融合并编码到每个视图的特征中, 有利于获取更好的光场超分辨率图像。

2.3 基于 SRFD 模块的数据重构

数据重构一般包括特征融合和特征上采样两部分, 这两部分逐步将网络提取的有效信息解码成最终的高清图像。本文在数据重构过程中, 借鉴了残差特征蒸馏的思想完成特征融合。残差特征蒸馏 (Residual feature distillation, RFD) 模块^[12] 的结构如图 6(a) 所示, 第一级平行的浅层残差块和 1×1 卷积块将输入特征分成两部分, 一部分保留, 另一部分馈送到下一层的浅层残差块和 1×1 卷积进行新的蒸馏操作, 最后将各级蒸馏特征进行通道拼接, 采用对比度感知的通道注意力 (contrast-aware channel attention, CCA) 机制完成有效特征提取。考虑到 RFD 模块常用于特征提取部分, CCA 层强化多通道有效特征学习, 而本文特征融合部分更侧重于前期学习特征的解码而不是特征学习, 故省去了 RFD 模块中的 CCA 层, 形成了 SRFD 模块。在 SRFD 模块中, 最前端增加了 3×3 卷积和 ReLU 激活来对输入特征进行细化学习, 同时在浅层残差块中将激活函数 ReLU 直接放到了卷积之后, 使得当前层的特征不经激活直接馈入下一蒸馏环节, 浅层信息和深层信息得到了更直接的结合。

数据经过融合操作后, 需将得到的重构特征馈送到上采样模块来实现光场图像的超分辨率。本文的上采样部分主要由像素洗牌操作完成^[4-7], 融合数据先通过 1×1 卷积将深度扩展到 $\alpha^2 C$ (其中 α 为超分系数, C 代表通道数), 然后进行像素洗牌操作, 将重构特征

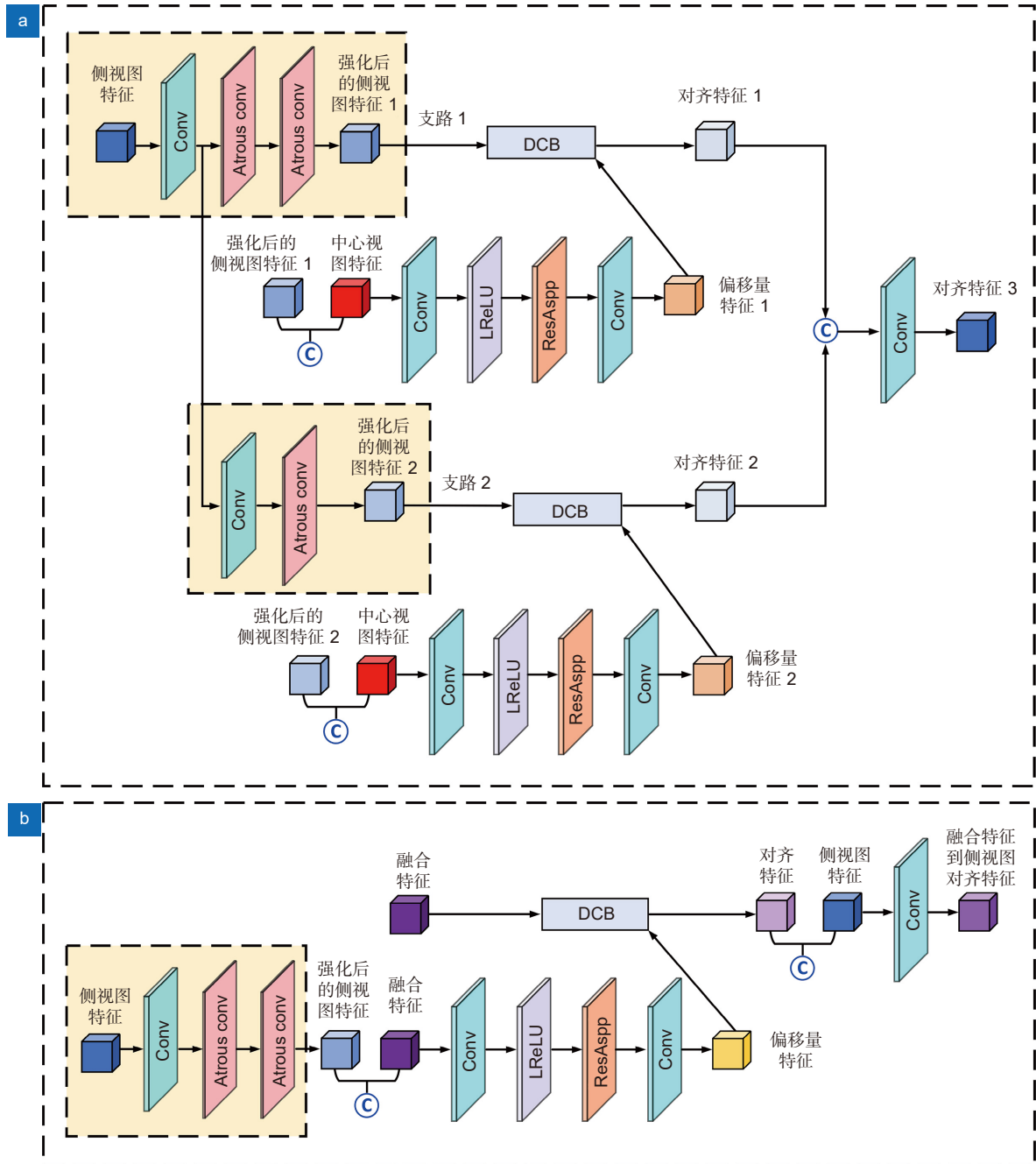


图 5 EADA 模块细节。(a) EADA 特征收集细节; (b) EADA 特征发散细节

Fig. 5 EADA module details. (a) EADA feature collection details; (b) EADA feature distribution details

的目标分辨率提升到 $\alpha S \times \alpha T$, 最后进行 1×1 卷积将特征通道数压缩到 1, 生成超分辨率后的子孔径图像阵列。

3 实验结果与分析

本文使用 5 个公共光场数据集 (EPFL^[13], HCInew^[14], HCIOld^[15], INRIA^[16] 和 STFgantry^[17]) 来进行训练和评估本文方法的有效性, 训练集和测试集的场景个数如

表 1 所示。先将原始的光场图像使用双三次插值下采样生成低分辨率光场图像, 通过本文搭建的光场超分辨率重构网络进行学习, 生成高分辨率的光场图像, 再将得到的高分辨率光场图像与原始光场图像进行对比。在训练阶段, 对每个子孔径图像裁剪, 并使用双三次下采样方法生成大小为 32×32 的低分辨率块, 通过数据增强作为训练数据集。网络使用 L1 损失函数进行训练, 并使用 Adam 方法^[18]进行优化。用零值初

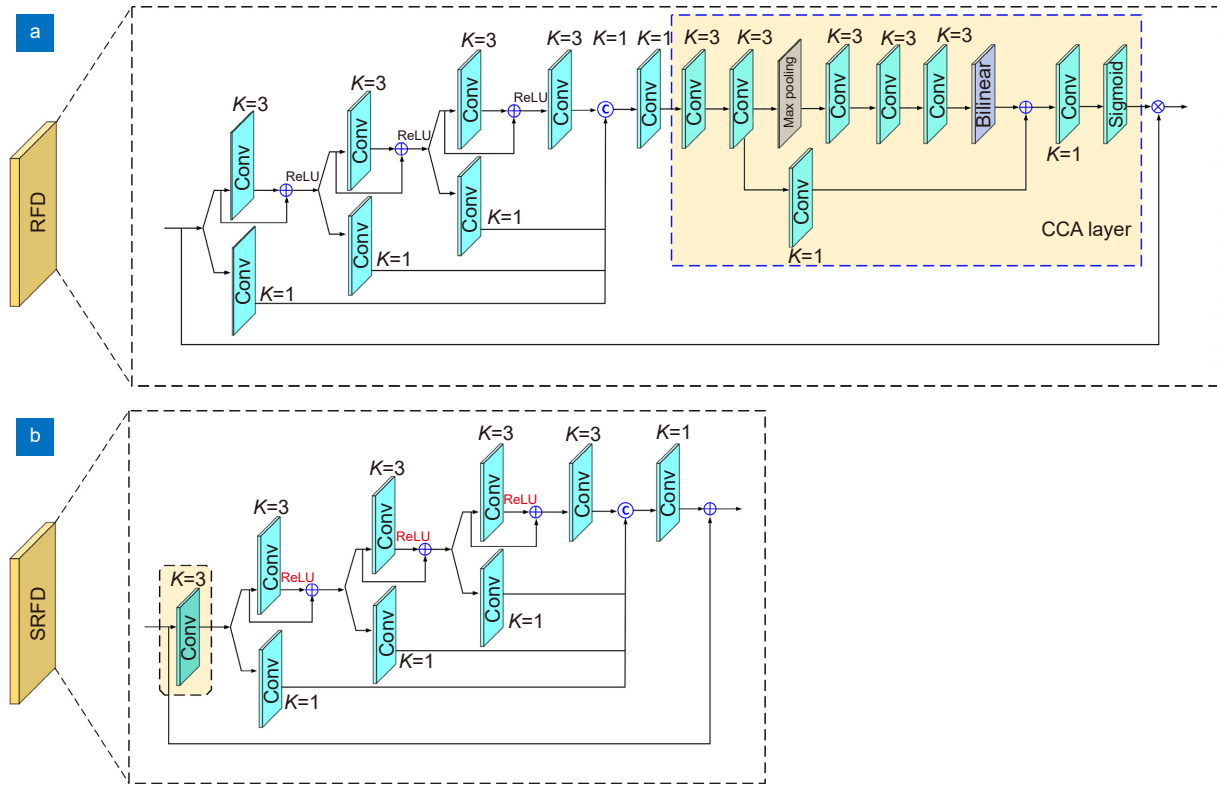


图 6 RFD 模块的简化。(a) RFD 模块细节; (b) SRFD 模块细节
Fig. 6 RFD module simplification. (a) RFD module details; (b) SRFD module details

表 1 实验使用的 5 个公共光场数据集

Table 1 Five public light field datasets used in our experiment

数据集	EPFL ^[13]	HCInew ^[14]	HCInold ^[15]	INRIA ^[16]	STFGantry ^[17]	总共
#训练	70	20	10	35	9	144
#测试	10	4	2	5	2	23

始化偏移量生成分支中最后一个卷积层的权重和偏差, 并使用 Kaiming 方法^[19] 初始化网络的其它部分。本文所有实验均在 Pytorch 平台下使用 Python 语言进行, 使用带有 NVIDIA GeForce RTX 3080 GPU 的图像处理设备对网络进行训练和测试。批尺寸设置为 2, 初始学习率设置为 2×10^{-4} , 每 15 个周期学习率降低一半, 训练在第 50 个周期后结束。

3.1 网络各模块有效性验证

3.1.1 MBR 块的有效性

为了证明 MBR 块中多支路连接的有效性, 在深

层特征提取模块、数据重构模块保持不变的基础上, 将 MBR 块与省略支路 2、支路 3 跳跃连接的 MBR 块分别作为浅层特征提取模块, 进行了对比实验。实验过程中对光场图像进行 4 倍超分辨率重构, 基于不同浅层特征提取模块的网络得到的光场超分子孔径图像的平均 PSNR、SSIM 值如表 2 所示。对比表 2 的数据可以看出, MBR 块使用了跳跃连接提取初级特征, 获得了更丰富的上下文信息, 使得最终的网络模型性能明显更好, 验证了本文设计浅层特征提取模块的有效性。

表 2 基于不同浅层特征提取模块的光场图像 4 倍超分 PSNR/SSIM 值

Table 2 PSNR/SSIM values achieved by different shallow feature extraction modules for 4×SR

模型	EPFL ^[13]	HCInew ^[14]	HCInold ^[15]	INRIA ^[16]	STFGantry ^[17]
无残差	25.26/0.8324	27.71/0.8517	32.58/0.9344	26.95/0.8867	26.09/0.8452
MBR	28.81/0.9190	31.30/0.9206	37.39/0.9725	30.81/0.9513	31.29/0.9511

3.1.2 EADA 模块的有效性

为了验证所设计用于深层特征提取的 EADA 模块的有效性, 在浅层特征提取模块、数据重构模块保持不变的基础上, 将 EADA 模块与 ADA 模块、单支路 EADA 模块分别作为深层特征提取模块, 进行了对比实验。单支路 EADA 模块是 EADA 模块对支路 2 的省略。实验中仍然对光场图像进行 4 倍超分辨率重构, 基于不同深层特征提取模块的网络得到的光场超分子孔径图像的平均 PSNR、SSIM 值如表 3 所示。对比表 3 的数据可以看出, EADA 模块充分挖掘了各个侧视图的特征信息, 很好地将角度信息融合并编码到每个视图的特征中, 获取了更好的光场超分辨率图像。因此, 验证了本文提出的深层特征提取模块的有效性。

3.1.3 SRFD 模块的有效性

为了验证所设计用于特征融合部分的 SRFD 模块的有效性, 在浅层特征提取模块、深层特征提取模块保持不变的基础上, 将 RFD 模块与 SRFD 模块分别作为特征融合模块进行了对比实验。实验中仍然对光场图像进行 4 倍超分辨率重构, 基于不同特征融合模块的网络得到的光场超分子孔径图像的平均 PSNR、SSIM 值如表 4 所示。对比表 4 的数据可以看出, 对 RFD 模块进行简化后的 SRFD 模块并没有降低超分后光场图像的 PSNR、SSIM 值, 说明本文采用的 SRFD 模块在保证特征融合效果的前提下, 有效节约了计算成本。

3.2 网络性能分析

为了对本文所提光场超分辨率重构网络性能进行评价, 本文对 5 个公共光场数据集的 23 个场景进行 2 倍超分和 4 倍超分, 分别从超分后图像主观视觉效

果展示、PSNR 和 SSIM 值计算、算法复杂度分析等三个方面对本文算法进行了定性分析与定量评价, 并与其它图像超分辨率重构方法进行了对比, 以验证本文算法的优越性。

3.2.1 主观视觉效果对比

为了定性分析本文超分辨率重构网络的性能, 将本文提出的算法与当下较为先进的两种单图像超分辨率方法 (EDSR^[20] 和 RCAN^[21]) 和四种光场图像超分辨率方法 (ResLF^[8]、LFSSR^[7]、LF-InterNet^[4] 和 LF-DFnet^[5]) 得到的超分图像进行了视觉效果对比, 如图 7~图 10 所示。

图 7 和图 8 展示的是 HCInew^[14] 数据集的 “Origami” 和 “Herbs” 场景的 2 倍超分效果, 在图 7 红色实线矩形框所框圆形桶的水平条纹 (椭圆形虚线框所框区域) 和三角形花纹区域 (矩形虚线框所框区域), 及图 8 红色实线矩形框所框两簇植物 (椭圆形及矩形虚线框所框区域), ResLF、LFSSR、LF-InterNet、LF-DFnet、本文算法的超分效果明显高于 EDSR 和 RCAN 算法的超分效果, 且在这些区域本文算法的超分效果较 ResLF、LFSSR、LF-InterNet 和 LF-DFnet 算法有轻微改善。图 9 和图 10 展示的是 INRIA^[16] 数据集的 “Bee” 场景和 STFgantry^[17] 数据集中 “Lego Knights” 场景的 4 倍超分效果, 在图 9 蓝色实线矩形框所框叶子的茎 (矩形虚线框所框区域), LF-InterNet 和本文算法的超分效果高于其它算法; 在图 10 蓝色实线矩形框所框小人的盾牌 (椭圆形及矩形虚线框所框区域), LF-DFnet、本文算法的超分效果明显高于其它算法。基于以上视觉效果对比, 说明本文所提超分算法能够获得较清晰的光场高分辨率图像。

分析其原因, EDSR 和 RCAN 算法仅超分光场中

表 3 基于不同深层特征提取模块的光场图像 4 倍超分 PSNR/SSIM 值

Table 3 PSNR/SSIM values achieved by different deep feature extraction modules for 4* SR

模型	EPFL ^[13]	HCInew ^[14]	HCInew ^[15]	INRIA ^[16]	STFgantry ^[17]
ADA模块	28.77/0.9172	31.26/0.9198	37.41/0.9723	30.80/0.9507	31.17/0.9497
单支路EADA模块	28.78/0.9166	31.21/0.9186	37.31/0.9717	30.85/0.9502	31.03/0.9474
EADA模块	28.81/0.9190	31.30/0.9206	37.39/0.9725	30.81/0.9513	31.29/0.9511

表 4 基于不同特征融合模块的光场图像 4 倍超分 PSNR/SSIM 值

Table 4 PSNR/SSIM values achieved by different feature fusion modules for 4* SR

模型	EPFL ^[13]	HCInew ^[14]	HCInew ^[15]	INRIA ^[16]	STFgantry ^[17]
RFD模块	29.01/0.9183	31.32/0.9198	37.39/0.9718	31.08/0.9509	31.14/0.9499
SRFD模块	28.81/0.9190	31.30/0.9206	37.39/0.9725	30.81/0.9513	31.29/0.9511

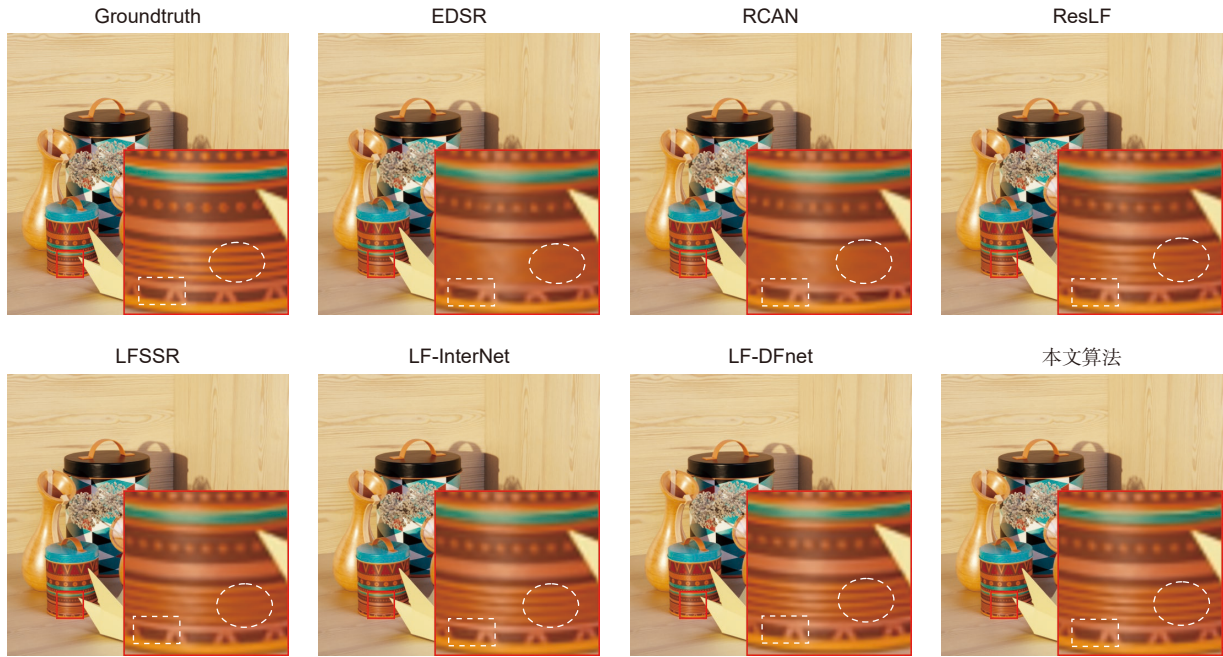


图 7 “Origami” 场景 2 倍超分的视觉效果对比
Fig. 7 Visual contrast of the "Origami" scene with 2^x SR

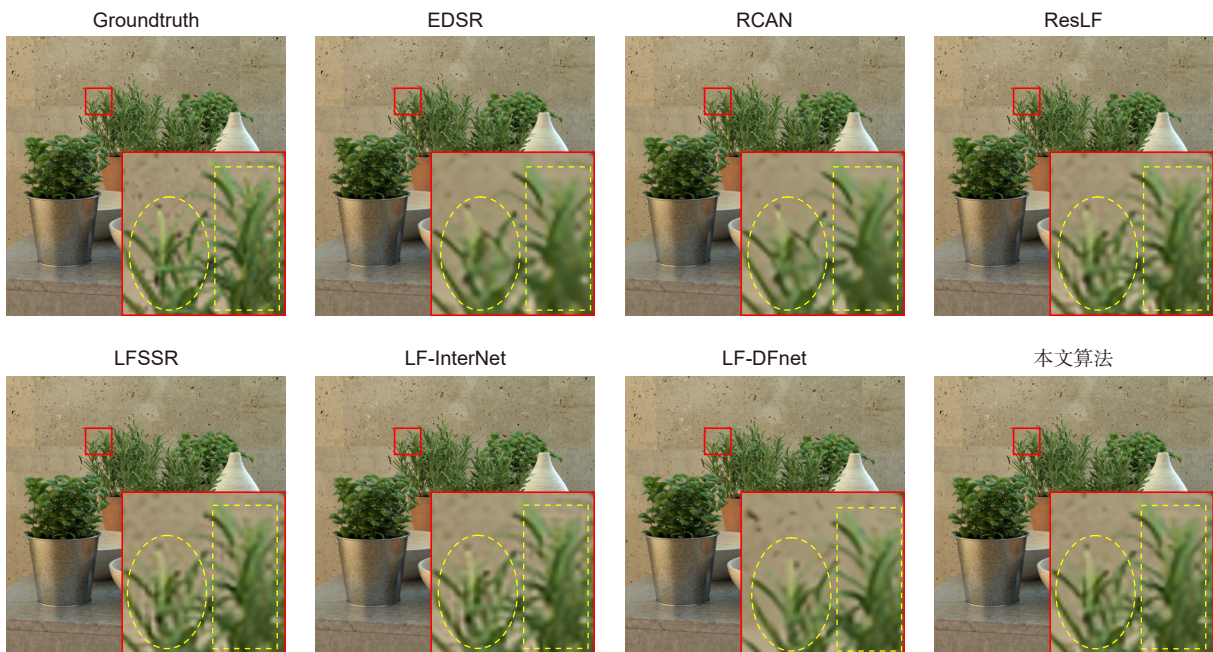


图 8 “Herbs” 场景 2 倍超分的视觉效果对比
Fig. 8 Visual contrast of the "Herbs" scene with 2^x SR

的单个视图, 并没有使用光场的 4D 信息, 网络超分性能最差; ResLF、LFSSR、LF-InterNet 算法侧重于独立学习空间特征和角度特征, 没有挖掘不同视图间的差异, 限制其超分性能; LF-DFnet 算法提出视差机制, 充分挖掘了不同视图间的视差信息, 超分性能

有所提升, 本文算法在挖掘不同视图间的视差信息前, 对侧视图自身特征进行了强化学习, 因此超分性能表现最好。

3.2.2 定量评价结果对比

定性评价受人眼视觉感知特性限制, 不易客观的

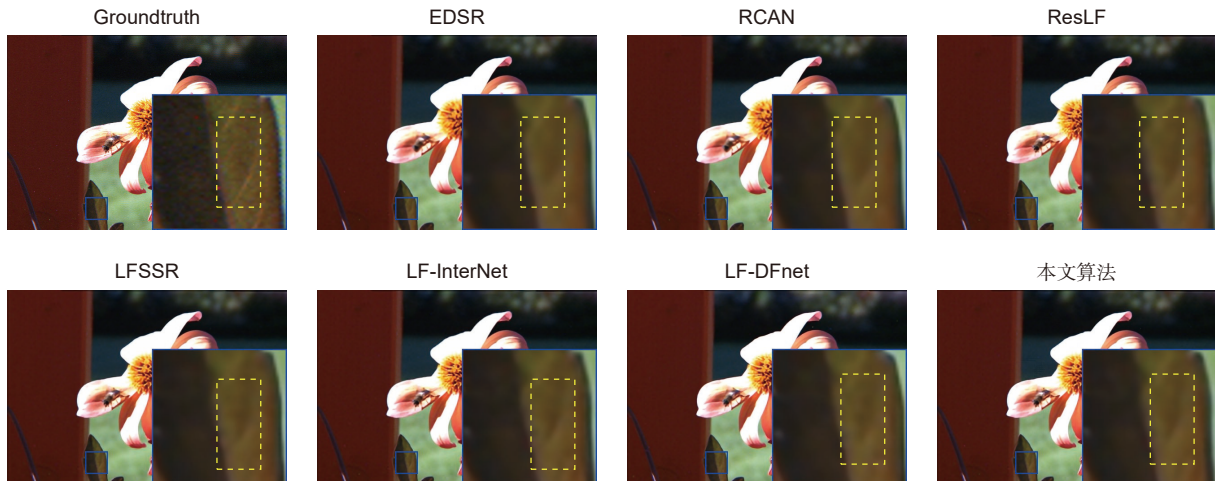


图 9 “Bee” 场景 4 倍超分的视觉效果对比
 Fig. 9 Visual contrast of the "Bee" scene with 4× SR

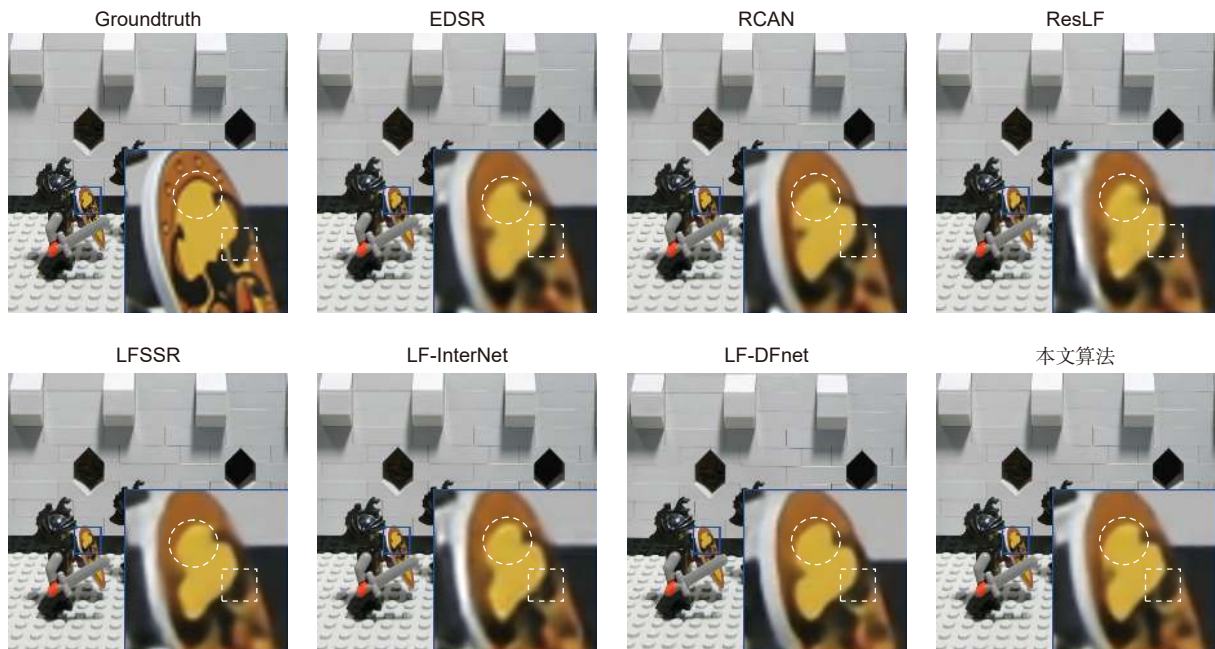


图 10 “Lego Knights” 场景 4 倍超分的视觉效果对比
 Fig. 10 Visual contrast of the "Lego Knights" scene with 4× SR

反映图像质量。为了进一步准确评价本文所提算法的超分辨率性能, 分别计算了 5 个公共光场数据集中每一测试场景 2 倍超分和 4 倍超分后的 PSNR 值和 SSIM 值, 并针对不同数据集取了平均值, 最后计算了所有数据集的均值, 不同算法的实验结果如表 5、表 6 所示。表中的蓝色加粗数字表示当前列的最大值, 黑色加粗字体表示当前列的次大值, 其所在位置分别代表了对应客观评价指标下的最优算法和次优算法。从表 5、表 6 可以看出, 除了在 INRIA 数据集 4 倍超

分的 PSNR 值是次大值外, 本文算法获取的高分辨率光场图像的 PSNR、SSIM 值均优于其它算法, 进一步验证了本文所提光场超分辨率重构网络算法的有效性。

3.2.3 算法复杂度

在验证了本文所提算法对光场图像的超分效果后, 进一步对算法的复杂度进行了分析, 分别对比了本文算法与其它几种算法的参数量 (Parameters)、每秒浮点运算次数 (Flo/s), 具体数据如表 7 所示。由表 7 可

表 5 不同算法对光场图像 2 倍超分 PSNR/SSIM 值

Table 5 PSNR/SSIM values achieved by different methods for 2^x SR

超分方法	EPFL ^[13]	HCInew ^[14]	HCInold ^[15]	INRIA ^[16]	STFgantry ^[17]	Average
EDSR ^[20]	33.09/0.9631	34.83/0.9594	41.01/0.9875	34.97/0.9765	36.29/0.9819	36.04/0.9728
RCAN ^[21]	33.16/0.9635	34.98/0.9602	41.05/0.9875	35.01/0.9769	36.33/0.9825	36.11/0.9741
ResLF ^[6]	32.75/0.9672	36.07/0.9715	42.61/0.9922	34.57/0.9784	36.89/0.9873	36.58/0.9793
LFSSR ^[7]	33.69/0.9748	36.86/0.9753	43.75/0.9939	35.27/0.9834	38.07/0.9902	37.53/0.9835
LF-InterNet ^[4]	34.14/0.9761	37.28/0.9769	44.45/0.9945	35.80/0.9846	38.72/0.9916	38.08/0.9847
LF-DFnet ^[5]	34.44/0.9766	37.44/0.9786	44.23/0.9943	36.36/0.9841	39.61/0.9935	38.42/0.9854
本文方法	34.58/0.9772	37.92/0.9796	44.84/0.9948	36.59/0.9854	40.11/0.9939	38.81/0.9862

表 6 不同算法对光场图像 4 倍超分 PSNR/SSIM 值

Table 6 PSNR/SSIM values achieved by different methods for 4^x SR

超分方法	EPFL ^[13]	HCInew ^[14]	HCInold ^[15]	INRIA ^[16]	STFgantry ^[17]	Average
EDSR ^[20]	27.84/0.8858	29.60/0.8874	35.18/0.9538	29.66/0.9259	28.70/0.9075	30.20/0.9121
RCAN ^[21]	27.88/0.8863	29.63/0.8880	35.20/0.9540	29.76/0.9273	28.90/0.9110	30.27/0.9133
ResLF ^[6]	27.46/0.8899	29.92/0.9011	36.12/0.9651	29.64/0.9339	28.99/0.9214	30.43/0.9223
LFSSR ^[7]	28.27/0.9080	30.72/0.9124	36.70/0.9690	30.31/0.9446	30.15/0.9385	31.23/0.9345
LF-InterNet ^[4]	28.67/0.9143	30.98/0.9165	37.11/0.9715	30.64/0.9486	30.53/0.9426	31.59/0.9387
LF-DFnet ^[5]	28.77/0.9165	31.23/0.9196	37.32/0.9718	30.83/0.9503	31.15/0.9494	31.86/0.9415
本文方法	28.81/0.9190	31.30/0.9206	37.39/0.9725	30.81/0.9513	31.29/0.9511	31.92/0.9429

表 7 不同算法 (2 倍超分/4 倍超分) 复杂度对比

Table 7 Comparisons of the number of parameters and FLOPs by different methods for 2^x SR and 4^x SR

Method	Parameters/M	FLOPs/G
EDSR ^[20]	38.62/38.89	39.56×25/40.66×25
RCAN ^[21]	15.31/15.36	15.59×25/15.65×25
ResLF ^[6]	6.35/6.79	37.06/39.70
LFSSR ^[7]	0.81/1.61	25.70/128.44
LF-InterNet ^[4]	4.80/5.23	47.46/50.10
LF-DFnet ^[5]	3.94/3.99	57.22/57.31
本文方法	12.74/12.80	238.92/240.51

知, 本文算法的参数数量、每秒浮点运算次数低于 EDSR、RCAN 算法, 但高于 ResLF、LFSSR、LF-InterNet、LF-DFnet 算法。与性能优越的 LF-InterNet、LF-DFnet 算法相比, 本文算法的参数数量及每秒浮点运算次数大约增加了 3 倍, 在合理控制复杂度的前提下, 获得了较高的光场超分效果。

4 结论

本文提出了一种基于角度差异强化的光场图像超分网络, 以提高光场子孔径图像的空间分辨率。在浅层特征提取部分, 采用了多分支残差 (multi-branch

residual, MBR) 块来提取 4D 光场中的固有结构信息; 在深层特征提取部分, 设计了强化的角度可变形对准 (enhanced angular deformable alignment, EADA) 模块强化各侧视图自身特征的挖掘, 使得视图间角度的差异得到更好的学习, 并很好地将角度信息融合并编码到每个视图特征中; 在数据重构部分, 采用了结构更为简单的残差特征蒸馏 (simplified residual feature distillation, SRFD) 模块, 使得特征融合中浅层信息和深层信息可以更直接结合。消融实验验证了本文所提网络各个模块的有效性。与其它光场超分辨率重构算法相比, 本文算法在合理控制参数数量的情况下, 主观

视觉测试和定量评价均表现出较高性能。另外, 针对本文网络模型参数量较高的问题, 后期将进一步简化数据重构部分的残差特征蒸馏模块, 在保证超分性能的前提下进一步提高网络运算速度。

参考文献

- [1] Ng R, Levoy M, Brédif M, et al. Light field photography with a hand-held plenoptic camera[J]. *Comput Sci Tech Rep CSTR*, 2005, 2(11): 1–11.
- [2] Tao M W, Hadap S, Malik J, et al. Depth from combining defocus and correspondence using light-field Cameras[C]// *Proceedings of 2013 IEEE International Conference on Computer Vision*, Sydney, 2013: 673–680. <https://doi.org/10.1109/ICCV.2013.89>.
- [3] Levoy M, Hanrahan P. Light field rendering[C]// *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, New York, 1996: 31–42. <https://doi.org/10.1145/237170.237199>.
- [4] Wang Y Q, Wang L G, Yang J G, et al. Spatial-angular interaction for light field image super-resolution[C]// *Proceedings of the 16th European Conference on Computer Vision*, Glasgow, 2020: 290–308. https://doi.org/10.1007/978-3-030-58592-1_18.
- [5] Wang Y Q, Yang J G, Wang L G, et al. Light field image super-resolution using deformable convolution[J]. *IEEE Trans Image Process*, 2020, 30: 1057–1071.
- [6] Mo Y, Wang Y Q, Xiao C, et al. Dense dual-attention network for light field image super-resolution[J]. *IEEE Trans Circuits Syst Video Technol*, 2022, 32(7): 4431–4443.
- [7] Yeung H W F, Hou J H, Chen X M, et al. Light field spatial super-resolution using deep efficient spatial-angular separable convolution[J]. *IEEE Trans Image Process*, 2019, 28(5): 2319–2330.
- [8] Zhang S, Lin Y F, Sheng H. Residual networks for light field image super-resolution[C]// *Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, 2019: 11046–11055. <https://doi.org/10.1109/CVPR.2019.01130>.
- [9] Jin J, Hou J H, Chen J, et al. Light field spatial super-resolution via deep combinatorial geometry embedding and structural consistency regularization[C]// *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, 2020: 2260–2269. <https://doi.org/10.1109/CVPR42600.2020.00233>.
- [10] Liang Z Y, Wang Y Q, Wang L G, et al. Angular-flexible network for light field image super-resolution[J]. *Electron Lett*, 2021, 57(24): 921–924.
- [11] Zhao Y Y, Shi S X. Light-field image super-resolution based on multi-scale feature fusion[J]. *Opto-Electron Eng*, 2020, 47(12): 200007.
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): 200007.
- [12] Liu J, Tang J, Wu G S. Residual feature distillation network for lightweight image super-resolution[C]// *Proceedings of the European Conference on Computer Vision*, Glasgow, 2020: 41–55. https://doi.org/10.1007/978-3-030-67070-2_2.
- [13] Rerabek M, Ebrahimi T. New light field image dataset[C]// *Proceedings of the 8th International Conference on Quality of Multimedia Experience*, Lisbon, 2016.
- [14] Honauer K, Johannsen O, Kondermann D, et al. A dataset and evaluation methodology for depth estimation on 4D light fields[C]// *Proceedings of the 13th Asian Conference on Computer Vision*, Cham, 2016: 19–34. https://doi.org/10.1007/978-3-319-54187-7_2.
- [15] Wanner S, Meister S, Goldluecke B. Datasets and benchmarks for densely sampled 4D light fields[M]// Bronstein M, Favre J, Hormann K. *Vision, Modeling and Visualization*. Eurographics Association, 2013: 225–226. <https://doi.org/10.2312/PE.VMV.VMV13.225-226>.
- [16] Le Pendu M, Jiang X R, Guillemot C. Light field inpainting propagation via low rank matrix completion[J]. *IEEE Trans Image Process*, 2018, 27(4): 1981–1993.
- [17] Vaish V, Adams A. The (new) Stanford light field archive, computer graphics laboratory, Stanford University[EB/OL]. 2008. <http://lightfield.stanford.edu>.
- [18] Kingma D P, Ba J. Adam: a method for stochastic optimization[C]// *Proceedings of the 3rd International Conference on Learning Representations*, San Diego, 2015.
- [19] He K M, Zhang X Y, Ren S Q, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification[C]// *Proceedings of 2015 International Conference on Computer Vision*, Santiago, 2015: 1026–1034. <https://doi.org/10.1109/ICCV.2015.123>.
- [20] Lim B, Son S, Kim H, et al. Enhanced deep residual networks for single image super-resolution[C]// *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Honolulu, 2017: 136–144. <https://doi.org/10.1109/CVPRW.2017.151>.
- [21] Zhang Y L, Li K P, Li K, et al. Image super-resolution using very deep residual channel attention networks[C]// *Proceedings of the 15th European Conference on Computer Vision*, Munich, 2018: 294–310. https://doi.org/10.1007/978-3-030-01234-2_18.

作者简介



吕天琪(1996-), 女, 硕士研究生, 研究方向: 光场信息获取与处理。

E-mail: 2505632776@qq.com

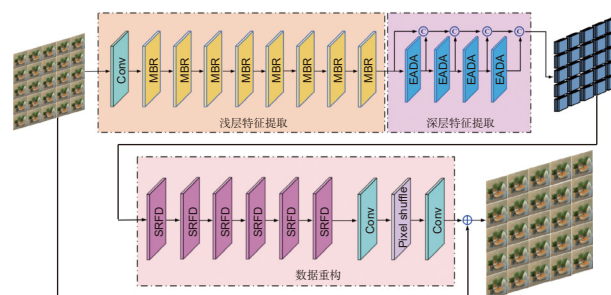


【通信作者】武迎春(1984-), 女, 博士, 教授, 研究方向: 3D 数据获取与处理、光场信息获取与处理。

E-mail: yingchunwu3030@foxmail.com

Light field image super-resolution network based on angular difference enhancement

Lv Tianqi, Wu Yingchun*, Zhao Xianling



Structure of light-field image super resolution network

Overview: Based on advanced imaging technology, light field cameras can obtain the spatial information and the angular information of the scene synchronously. It achieves higher dimensional scene representation by sacrificing spatial resolution. To improve the spatial resolution of the light field image, a light field super-resolution reconstruction network based on angle difference enhancement is built in this paper. In the proposed network, eight multi-branch residual blocks are used to extract shallow features. Then, four enhanced angular deformable alignment modules are used to extract deep features. Finally, six simplified residual feature distillation modules and pixel shuffle modules are used to complete data reconstruction. The proposed network takes advantage of the angle difference of the light field to complete the spatial information super-resolution. To obtain more features difference between different views, the own feature of the single view is emphasized during the feature extraction. The design ideas of different modules are as follows.

a) Shallow feature extraction based on multi-branch residual (MBR) block. The convolution blocks of 1×1 are used to complete the channel map first, then eight multi-branch residual blocks are used to extract shallow features. By using the skip connection and atrous convolution in the multi-branch residual blocks, the network degradation is suppressed, more hierarchical information is extracted and the receptive field of the feature is expanded.

b) Deep feature extraction based on enhanced angular deformable alignment (EADA) block. To take full advantage of the angle information correlation of multi-view images in the 4D light field, the ADA block is replaced by the EADA block to complete the feature collection and feature distribution. Compared with the ADA block, EADA adds a branch in front of the two DCB blocks in the feature collection. Two atrous convolutions are added in branch 1, and one ordinary convolution and one atrous convolution are added in branch 2. In the feature distribution, one ordinary convolution and two atrous convolutions are added before DCB blocks. By using the EADA block, the own feature of each side-view is fully extracted, and the more accurate disparity is obtained. So the super-resolution performance of the network can be improved.

c) Data reconstruction based on the simplified residual feature distillation (SRFD) block. The RFD block is adopted in the proposed light field super-resolution network to complete the data reconstruction. To reduce the parameters of the network, the RFD block is simplified to obtain SRFD the block by omitting the contrast-aware channel attention (CCA) layer. A convolution and Relu activation are added at the front of the SRFD block, so the input feature is refined; In the shallow residual block, the activation function is swapped with convolution. In this way, the feature of the current layer is fed directly into the next distillation block without activation, which makes the shallow information and the deep information get a more direct combination.

The experiments are carried out from the aspects of validity verification of each network module, comparison of subjective visual effects, comparison of quantitative evaluation results, and algorithm complexity. The performance of the proposed network is verified on five public light field data sets. The proposed algorithm obtains high-resolution light field sub-aperture images with higher PSNR and SSIM.

Lv T Q, Wu Y C, Zhao X L. Light field image super-resolution network based on angular difference enhancement[J]. *Opto-Electron Eng*, 2023, 50(2): 220185; DOI: 10.12086/oe.2023.220185

Foundation item: National Natural Science Foundation of China (61601318), the Basic Research Project of Shanxi Province (202103021224278), and Research Project Supported by Shanxi Scholarship Council of China (2020-128)

School of Electronic and Information Engineering, Taiyuan University of Science and Technology, Taiyuan, Shanxi 030024, China

* E-mail: yingchunwu3030@foxmail.com