

文章编号:1007-2780(2022)10-1345-10

基于空间角度解耦融合网络的光场重建

张洪基, 邓慧萍*, 向 森, 吴 谨

(武汉科技大学 信息科学与工程学院, 湖北 武汉 430081)

摘要:密集采样的光场能提供沉浸式的视觉体验,在三维重建和虚拟现实等领域有着广泛的应用。光场重建技术可以从稀疏采样光场获取密集采样光场,但当光场图像的基线增大时,会造成重建视图中边缘的遮挡和纹理重复区域的模糊。本文提出了一种基于空间角度解耦融合的光场重建网络,充分利用光场图像蕴含的深度信息以及空间信息解决上述复杂区域的重建难题。采用多路输入方式来获取光场多方向信息以解决遮挡问题,在每条支路通过空洞空间卷积池化金字塔模块来提取多尺度特征,以获取大接受域的视差信息和上下文信息;设计了空间角度解耦融合模块,利用空间信息指导光场角度超分辨率重建,使重建视图在纹理重复区域更加清晰。实验结果表明,该网络在 HCI、HCI old 和 30 scenes 数据集上重建光场的平均 PSNR 和 SSIM 分别达到了 39.73 dB 和 0.965,表现出较高的准确性,优于所比较的方法。

关键词:光场图像;角度超分辨率;深度学习;解耦融合

中图分类号:TP391 文献标识码:A doi:10.37188/CJLCD.2022-0117

Light field reconstruction based on spatial-angular decouple and fuse network

ZHANG Hong-ji, DENG Hui-ping*, XIANG Sen, WU Jin

(School of Information Science and Engineering, Wuhan University of Science and Technology,
Wuhan 430081, China)

Abstract: The densely sampled light field can provide immersive visual experience and have extensive applications in areas such as 3D reconstruction and virtual reality. Light field reconstruction technology can obtain densely sampled light fields from sparsely sampled light fields, but when the baseline of the light field image increases, it will cause occlusion on edges and textured duplicate areas blur in the reconstructed view. In this paper, we propose a light field reconstruction network based on spatial-angular decouple and fuse, it can make full use of the depth information and spatial information contained in the light field image and solve the reconstruction problems of the complex regions. The multi-branch input is used to supply the multi-directional information of the light field to solve the occlusion problem. In each branch, multi-scale features are extracted by atrous spatial pyramid pooling module to obtain disparity information and context information for the large receptive fields. The spatial-angular decouple and fuse module is designed to explore the spatial information, guiding the light field angle super-resolution reconstruction so as to obtain

收稿日期:2022-04-11;修订日期:2022-04-19.

基金项目:国家自然科学基金(No. 61702384, No. 61502357)

Supported by National Natural Science Foundation of China (No. 61702384, No. 61502357)

*通信联系人, E-mail: denghuiping@wust.edu.cn

the reconstruction view with clearer texture duplicate region. Experimental results show that the average PSNR and SSIM reached 39.73 dB and 0.965 on the HCI, HCI old and 30 scenes datasets, showing high accuracy and outperforming the compared method.

Key words: light field image; angular super resolution; deep learning; decouple and fuse

1 引言

光场相机可以在一次摄影曝光中从多个视角捕捉一个场景,记录场景的二维空间信息和二维角度信息,在深度估计、三维重建等领域具有重要的研究价值^[1-2]。然而,由于光场相机内部传感器分辨率有限,捕获的光场图像在满足空间分辨率需求时只能在角度域稀疏采样。因此,通过稀疏采样光场合成中间视图的光场重建^[3-5]成为光场应用的关键技术。光场角度分辨率重建又称为视点合成或视点绘制。依据是否依赖深度图的辅助可分为直接三维重建和基于深度辅助的光场重建。

传统的三维重建常采用信号的频域分析方法。Shi 等人^[6]只对边界或对角线子孔径图像进行采样,利用傅里叶域的稀疏性分析来恢复全光场,然而这种方法需要以特定的模式捕获光场,这限制了它的应用。Vagharshakyan 等人^[7]对光场极平面图像(Epipolar Plane Image, EPI)进行频域分析,在傅里叶域中采用了一种自适应的离散剪切变换来消除引入混叠效应^[8]的高频光谱。Pendurkar 等人^[9]新提出了一种称为傅里叶视差层的光场表示方法,通过移动不同的视差层对光场进行重建。以上从频域重建的方法在一定程度上能获取密集采样的光场,但是针对深度层次较多的复杂场景时,光场的频域特点难以分析,频谱更加复杂,频域混叠问题难以解决。

基于深度学习的三维重建方法,利用卷积神经网络强大的学习能力,通过堆叠的卷积层将稀疏采样光场映射到密集采样光场。Yeung 等人^[10]将整个四维光场输入一个伪四维卷积网络,并提出了一种空间-角度交替卷积来迭代细化重建结果。Meng 等人^[11-12]使用密集连接的4D卷积对光场进行上采样和优化,实现了在空间或角度维度的多个尺度上的重建。虽然4D卷积可以同时提取光场的空间和角度信息,但是会导致过拟合并且计算代价非常高。以上利用深度学习的

直接重建方法虽然无需对频谱进行分析,但是以稀疏采样的光场,特别是大基线光场作为输入时,提供的信息十分有限,尽管使用了高维卷积来提取特征,仍然难以建立起稀疏输入和密集光场之间的关系,需要复杂的优化网络进行后处理。

基于深度辅助的光场重建由于增加了光场的深度信息,往往能取得更好的重建质量。经典的3D warping算法^[13-14]就是通过深度信息的辅助,使参考视点与待合成视点建立几何映射关系,从而合成新视点。目前,也有不少工作将深度信息利用到基于深度学习的光场重建框架中。Flynn 等人^[15]利用颜色、深度和纹理先验,首次实现了使用神经网络来预测新视图。Kalantari 等人^[16]采用顺序卷积网络估计深度和颜色,然后使用连续的二维空间卷积层融合从不同输入视点映射到新视点的视图来产生最终的重建。以上早期的重建网络考虑到了使用颜色等先验信息辅助重建,但是颜色等信息只是光场空域特征的一部分,空间信息未能得到充分利用。Jin 等人^[17]首先估计出新视点的深度图,然后通过空间角度交替细化进行重构。该网络在细节方面能保持较好的效果,但由于是将所有子孔径图像一同输入顺序连接的卷积层来估计深度,忽略了光场视图之间的方向信息,导致了严重的遮挡问题。

除了显式的深度估计以外,Wu 等人^[18]利用编解码网络对不同剪切值的EPI进行评估,网络隐式地使用深度信息来选择重建良好的EPI,然后使用金字塔分解重建技术融合多个剪切EPI来得到密集采样EPI。该方法不需要显式地估计出深度图,但是EPI只是四维光场的二维采样,使用EPI进行重建没有完全利用到光场的空间角度信息。Zhou 等人^[19]提出了一种新的基于学习的多平面图表示方法,通过对不同图像的alpha通道进行融合来合成新的视图。Mildenhall 等人^[20]进一步提出利用多平面图来合成一个局部光场。基于多平面图的方法对小基线视图有着不错的重建效

果,但这种未显式估计出深度图的方法由于受深度平面数量的限制,在面对大基线输入时会超出极限外推范围,使重建视图边缘产生伪影。

综上所述,利用深度辅助的光场重建算法由于引入了额外的深度信息,有利于提高大基线条件下稀疏光场的角度超分辨率重建质量,但是现有的基于深度的方法还存在几个方面的问题与挑战。首先,视点间基线增大会使深度估计面临严重的遮挡问题,遮挡区域的深度估计错误将会导致重建视图边缘模糊;其次,现有的方法仅考虑了光场角维度的重建,没有充分利用到光场空域信息,未能充分挖掘光场蕴含的四维特征,使重建视图在纹理重复区域发生错误。因此,本文从提高深度估计的准确性并充分利用空域信息的角度出发,提出了一种基于空间角度解耦融合的光场重建网络。该网络主要包含一个多路输入的深度估计模块和一个整合光场空间角度信息的光场融合模块,在获得较为精确的重建视图的同时,保留了更多的纹理细节。该深度学习网络具有以下特点:

(1)充分利用光场的多方向信息,采用多支路输入方式,更好地解决遮挡问题;对每条支路使用空洞空间卷积池化金字塔模块(Atrous Spatial Pyramid Pooling, ASPP)以引入多尺度信息,通过增大感受野来捕捉更多的视差信息并获取丰富的上下文信息,提高深度估计的准确性。

(2)设计了空间角度解耦融合模块(Spatial-Angular Decouple and Fuse Module, SADFM),提取了光场空间维度蕴含的纹理信息和角维度

蕴含的视差信息并使之融合。在模块内使用残差连接方式将浅层特征与融合之后的深层特征连接,获取更丰富的特征表示。引入空间信息辅助角度超分辨率,使重建结果在纹理重复区域更加清晰。

2 光场重建算法

2.1 整体网络结构

一个4D光场可以表示为 $L(x, u)$,其中 $x=(x, y)$ 是空间坐标, $u=(u, v)$ 是角度坐标。光场角度超分辨率可以解释为:

$$\hat{L}(x, \hat{u}) = f(L(x, u)), \quad (1)$$

其中 $L(x, u)$ 为输入稀疏采样的光场, $\hat{L}(x, \hat{u})$ 为密集采样的光场, $\hat{u}=(\hat{u}, \hat{v})$ 为密集光场的角度坐标, f 表示要学习的重建过程的函数。

重建过程的整体框架如图1所示。本文利用了具有多个特征提取支路的深度估计模块来获取密集采样光场的深度图 $D(x, \hat{u})$,然后通过3D warping将输入视点 $L(x, u)$ 映射到深度图 $D(x, \hat{u})$ 上,映射得到的初始光场记为 $W(x, \hat{u}, u)$ 。再把 $W(x, \hat{u}, u)$ 输入空间角度解耦融合模块以充分探索光场空间角度之间的关系,最后通过3D卷积得到重建的残差图 $f_b(W(x, \hat{u}, u))$ 。

最终重建的密集采样光场可以表示为:

$$\hat{L}(x, \hat{u}) = W(x, \hat{u}, u_1) + f_b(W(x, \hat{u}, u)), \quad (2)$$

其中 $W(x, \hat{u}, u_1)$ 是从第一个输入视图 u_1 映射到深度图 $D(x, \hat{u})$ 的光场, f_b 表示光场视图融合过程的函数。

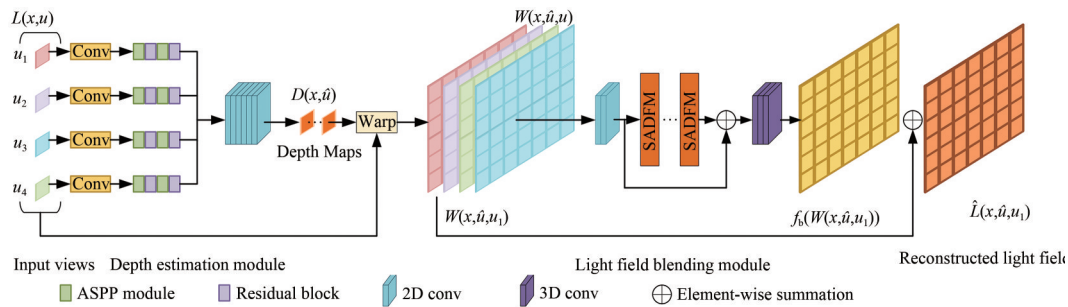


图1 本文的整体框架

Fig. 1 Overall framework of the proposed network

2.2 深度估计模块

深度估计模块利用输入的稀疏采样光场估计

出待重建密集采样光场各视点的深度,可以表示为:

$$D(x, \hat{u}) = f_d(L(x, u)), \quad (3)$$

其中 f_d 为表示深度估计过程的函数。

深度估计的准确性对光场重建的结果至关重要,但是在图像边缘等区域,特别是大基线输入视图的边缘,估计出的深度图往往具有严重的遮挡问题。本文采用多路输入形式,将每一个输入子孔径图像单独作为一个支路进行特征提取,各个支路之间权重共享。多路输入形式可以获得每个子孔径图像独特的位置信息和方向信息,在一定程度上解决了遮挡问题。

在每条支路上,为了获取大基线输入视图之间的视差信息,需要一个具有密集像素采样率的大接受域来提取层次特征。本文使用两个交替的 ASPP 模块和残差块作为特征提取模块。ASPP 模块如图 2 所示,首先对输入的每张子孔径图像进行 1×1 卷积,生成初始特征,然后输入 ASPP 模块进行深度特征提取,在每个 ASPP 块中,利用 3 个 3×3 空洞卷积(空洞率分别为 1、2、4)并行合并,引入了多尺度信息并增大感受野,提取到丰富的上下文信息和有密集采样率的层次特征。用参数为 0.1 的 leaky ReLU 层进行激活后,将这三个分支的特征通过 1×1 的卷积连接和融合。

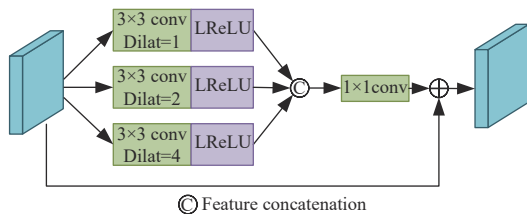


图 2 空间空洞金字塔模块

Fig. 2 ASPP module

在 4 张子孔径图分别经过特征提取之后,将提取的特征进行拼接并使用 6 个连续的卷积层来重建深度图,在加深网络的同时增大了感受野,进一步解决大基线问题。

2.3 视图 3D warping 模块

基于得到的密集采样光场各视图的深度图,可以通过 3D warping 建立稀疏采样光场和该深度图的映射关系,获取一系列从不同输入视图 warping 的光场图像,具体过程如式(4)所示:

$$W(x, \hat{u}, u) = L(x + D(x, \hat{u})(\hat{u} - u), u). \quad (4)$$

由于不同的子孔径图像拥有各自的视点信息,对于待合成的目标子孔径视图有着不同的视

角贡献,本文使用了从输入视点 warping 到其他待合成视点的所有视图进行融合。在框架中没有使用真实深度图,通过最小化 warping 光场的重建误差为深度估计网络提供适当的指导。此外,通过惩罚空间梯度来确保每个深度图的平滑性。深度估计模块的训练损失 l_d 如式(5)所示:

$$l_d = \sum_{x, \hat{u}} \left(\sum_u |L(x, \hat{u}) - W(x, \hat{u}, u)| + \nabla_x D(x, \hat{u}) \right), \quad (5)$$

其中 $L(x, \hat{u})$ 为真实光场图像。

2.4 光场融合模块

在 3D warping 之后,需要融合从不同输入视点 warping 的光场视图来得到残差图,然而,在 warping 过程中子孔径图像之间的视差不一致性将不可避免地造成高频损失。为了缓解这种高频损失,需要在角度重建过程中充分挖掘光场空域蕴含的二维图像信息。现有的融合方式未能利用完整的光场空间信息^[16],或者未能提取到光场空域的浅层特征^[17]。这些方法不能充分利用光场所蕴含的空间信息,造成重建视图纹理重复区域的模糊。因此,在重构光场的过程中除了要关注角度信息,还要注重光场的空域特性。

本文利用空间卷积和角度卷积来分别提取光场的空间和角度信息,并设计了 SADFM 来充分探索光场空间角度之间的关系。在重构过程中额外引入的空间信息可以增强空域细节,使重建视图的纹理重复区域更加清晰。SADFM 如图 3 所示,空间和角度卷积将特征张量分别变换为空间图像的堆栈 $X \times Y \times UV$ 和角块堆栈 $U \times V \times XY$,其中 X, Y 表示空间维度的大小, U, V 表示角维度的大小,然后实行相应的卷积操作。

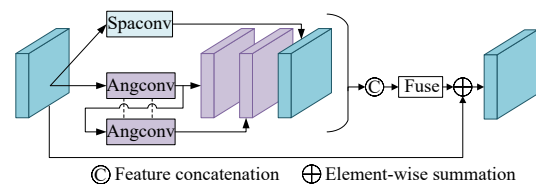


图 3 空间角度解耦融合模块

Fig. 3 Spatial-angular decouple and fuse module

卷积核的大小设置为 3×3 , 步长为 1, 在空间卷积上设置空洞率为 2 以增大空间感受野, 保证

在较大基线光场中仍能捕获细节特征。由于光场特征张量 $XY \times UV$ 中空间维度 XY 远大于角维度 UV , 在重建步骤中可利用的角度信息远小于空间信息, 因此本文利用了一次空间卷积和两个级联的角度卷积来分别提取空间信息和角度特征, 然后将所有特征进行连接并输入 1×1 卷积进行深度特征融合。级联的角度卷积权重共享, 不仅能减小模型大小, 还能加深网络, 提取更深层次的角度特征。然后使用残差连接方式将初始特征与融合之后的特征进行连接, 用浅层特征强化特征张量, 形成层次表示。

本文以残差连接方式级联 6 个 SADFM 加深网络, 在空间角度信息不断的解耦和融合步骤中整合信息流, 充分探索光场视图空间角度之间的关系, 进一步提高重建视图的质量。光场融合模块的损失函数为 l_b , 通过在最小化预测光场的重构误差来监督光场后续的重建:

$$l_b = \sum_{x, \hat{u}} |L(x, \hat{u}) - \hat{L}(x, \hat{u})|. \quad (6)$$

考虑到高质量的重建光场在 EPI 上应有严格的线型结构, 因此, 本文使用了一种基于 EPI 梯度的损失函数 l_e 来对输出 EPI 添加额外的约束^[17], 以进一步保持光场的视差结构, 增加新视点之间的视觉连续性。

$$l_e = \sum_{y, \hat{v}} \left(\left| \nabla_x E_{y, \hat{v}}(x, \hat{u}) - \nabla_x \hat{E}_{y, \hat{v}}(x, \hat{u}) \right| + \left| \nabla_{\hat{u}} E_{y, \hat{v}}(x, \hat{u}) - \nabla_{\hat{u}} \hat{E}_{y, \hat{v}}(x, \hat{u}) \right| \right) + \sum_{x, \hat{u}} \left(\left| \nabla_y E_{x, \hat{u}}(y, \hat{v}) - \nabla_y \hat{E}_{x, \hat{u}}(y, \hat{v}) \right| + \left| \nabla_{\hat{v}} E_{x, \hat{u}}(y, \hat{v}) - \nabla_{\hat{v}} \hat{E}_{x, \hat{u}}(y, \hat{v}) \right| \right), \quad (7)$$

其中, l_e 为重建 EPI($\hat{E}$) 和真实 EPI(E) 梯度之间的 l_1 距离, 梯度在水平和垂直 EPI 上沿空间和角度尺寸计算。

3 实验结果及分析

3.1 实验设置

本文提出的网络使用合成光场图像 HCI^[21]、HCI old^[22] 和 Lytro Illum 相机拍摄的真实光场图像 30 scenes^[16] 进行训练和测试。从以上数据集的 9×9 个视点中选取左上角的 7×7 个视点作为真实密集采样光场, 然后对其进行稀疏下采样,

获取 7×7 子孔径阵列的 4 个角落的视图作为稀疏输入的光场, 实验采用提出的方法从 2×2 稀疏采样光场重建出 7×7 的密集采样光场。

具体来说, 使用来自 HCI 数据集的 20 张合成图像进行训练, 使用 HCI、HCI old 中的测试集和来自 Lytro Illum 相机拍摄的 30 scenes 进行测试。这些数据集涵盖了评价光场重建方法的几个重要因素, 合成数据集包含高分辨率的纹理, 以测量保持高频细节的能力。真实数据集 30 scenes 包含了丰富的遮挡场景, 可以评估不同方法在自然照明和实际相机失真下对遮挡的处理能力。此外, 合成光场数据集包含视差范围为 $[-4, 4]$ 的大基线光场图像。在 7×7 的光场中, 视图之间的基线范围是视差的 6 倍, 即在 $[-24, 24]$ 的范围内, 远远大于商业相机捕获的光场的视差范围。这验证了本文提出方法在大基线输入的有效性。

所有用于训练的图像的空间分辨率为 512×512 。在训练过程中, 为了解决数据量不足的问题, 把每张图像随机裁剪成 96×96 的块进行训练。为了保持空间分辨率不变, 所有卷积层都使用了零填充。模型在 NVIDIA GTX3080Ti GPU 上运行, 采用 PyTorch 实现。实验使用了 Adam 优化器^[23] 进行优化。 β_1 和 β_2 分别设置为 0.9, 0.999。学习率最初设置为 $1e^{-4}$, 每 $5e^3$ 个周期降低 0.5 倍。Batch size 设为 1, 训练大约花了 6 天时间。最终的损失函数为 $\min(l_d + l_b + \lambda l_e)$ 。其中, λ 为 EPI 梯度损失的权重。

为了验证本文所提方法的有效性, 将本文的实验结果与一些利用深度信息的光场重建方法进行了定性和定量的比较。对比的方法有 Kallantari 等人^[16]、Wu 等人^[18]、Jin 等人^[17] 的方法。为了公平比较, 本文使用作者直接公布的光场重建结果或者使用它们公开的参数和代码进行了测试。

3.2 定量与定性分析

本文采用峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 和结构相似度指数 (Structural Similarity, SSIM) 来评估光场重建算法的性能。表 1、2、3 显示了在 HCI、HCI old 以及 30 scenes 上的定量比较。每个场景的最优值用加粗表示。

从表 1、2、3 中可以看出, 本文算法的 PSNR 和 SSIM 在除了 stillLife 的所有场景上均高于所比较的算法, 在 stillLife 场景中略低于 Jin 等人^[17]

的方法,处于次优值。在 HCI、HCI old 和 30 scenes 数据集上的平均 PSNR 比次优算法分别提高了 1.0, 0.68, 0.6 dB, 平均 SSIM 提高了 0.013, 0.001, 0.002。

表 1 不同方法在 HCI 数据集上的 PSNR(dB)/SSIM

Tab. 1 PSNR (dB)/SSIM of different methods over HCI datasets

Light field	Data type	Kalantari <i>et al.</i> ^[16]	Wu <i>et al.</i> ^[18]	Jin <i>et al.</i> ^[17]	Ours
bedroom	Synthetic	28.93/0.835	33.73/0.907	41.98/0.975	42.65/0.978
bicycle	Synthetic	24.65/0.795	29.94/0.878	34.03/0.954	35.51/0.966
dishes	Synthetic	22.87/0.740	24.19/0.770	32.76/0.882	32.84/0.894
herbs	Synthetic	24.84/0.665	28.36/0.755	29.63/0.938	31.38/0.960
Avg. over 4 light fields		25.32/0.759	29.05/0.828	34.60/0.937	35.60/0.950

表 2 不同方法在 HCI old 数据集上的 PSNR(dB)/SSIM

Tab. 2 PSNR (dB)/SSIM of different methods over HCI old datasets

Light field	Data type	Kalantari <i>et al.</i> ^[16]	Wu <i>et al.</i> ^[18]	Jin <i>et al.</i> ^[17]	Ours
buddha	Synthetic	29.70/0.908	40.92/0.980	45.65/0.991	47.16/0.993
buddha2	Synthetic	27.64/0.882	37.86/0.965	41.48/0.975	42.02/0.977
monasroom	Synthetic	28.89/0.902	41.09/0.984	45.88/0.990	46.71/0.992
papillon	Synthetic	28.66/0.887	41.98/0.980	45.51/0.987	46.19/0.989
stillLife	Synthetic	21.32/0.645	22.29/0.664	25.67/0.854	25.53/0.852
Avg. over 5 light fields		27.24/0.845	36.83/0.915	40.84/0.960	41.52/0.961

表 3 不同方法在 30 scenes 数据集上的 PSNR(dB)/SSIM

Tab. 3 PSNR (dB)/SSIM of different methods over 30 scenes datasets

Light field	Data type	Kalantari <i>et al.</i> ^[16]	Wu <i>et al.</i> ^[18]	Jin <i>et al.</i> ^[17]	Ours
Flower1	Real-world	23.17/0.844	33.97/0.970	37.42/0.983	37.95/0.987
Flower2	Real-world	22.48/0.811	33.17/0.961	36.52/0.981	36.94/0.983
IMG1324	Real-world	35.36/0.955	47.17/0.989	49.85/0.990	50.07/0.991
IMG1328	Real-world	31.85/0.892	41.93/0.976	44.99/0.985	45.57/0.986
IMG1528	Real-world	22.03/0.799	23.67/0.825	31.67/0.964	33.65/0.973
IMG1541	Real-world	23.11/0.757	24.94/0.830	35.56/0.973	36.78/0.980
Avg. over 30 light fields		27.81/0.858	38.10/0.961	41.47/0.983	42.07/0.985

实验在 HCI、HCI old 两个大基线合成数据集中选取了具有丰富纹理信息的 bicycle 和 buddha,在真实场景数据集 30 scenes 上选取了具有丰富遮挡的 IMG1528 和 IMG1555 来定性分析所提出算法的有效性。

在具有丰富高频特征的合成数据集上,对纹理重复区域的重建效果进行比较,如图 4 所示。由于引入了空间信息来指导角度重建,本文方法在 bicycle 的镂空金属网格区域和 buddha 的白色纹理区域都取得了较为清晰的重建效果。

Kalantari 等人^[16]和 Wu 等人^[18]未能充分利用

光场空间信息,纹理重复区域重建模糊;Jin 等人^[17]在重建过程中引入了空间卷积,因此在上述纹理重复区域重建出较为清晰的结果。但是他们的方 法未能使用浅层特征,在自行车(bicycle)的金属网格区域产生了比本文方法模糊的效果。

为了证明本文算法对遮挡的处理能力,在合成和真实世界数据集上对各个方法进行比较,如图 4、图 5 所示。Wu 等人^[18]隐式地利用了深度信息。受限于剪切值的数量,他们的方法不能应对复杂遮挡的区域,会产生严重的模糊和伪影,如 IMG1528 和 IMG1541;Kalantari 等人^[16]的方法虽

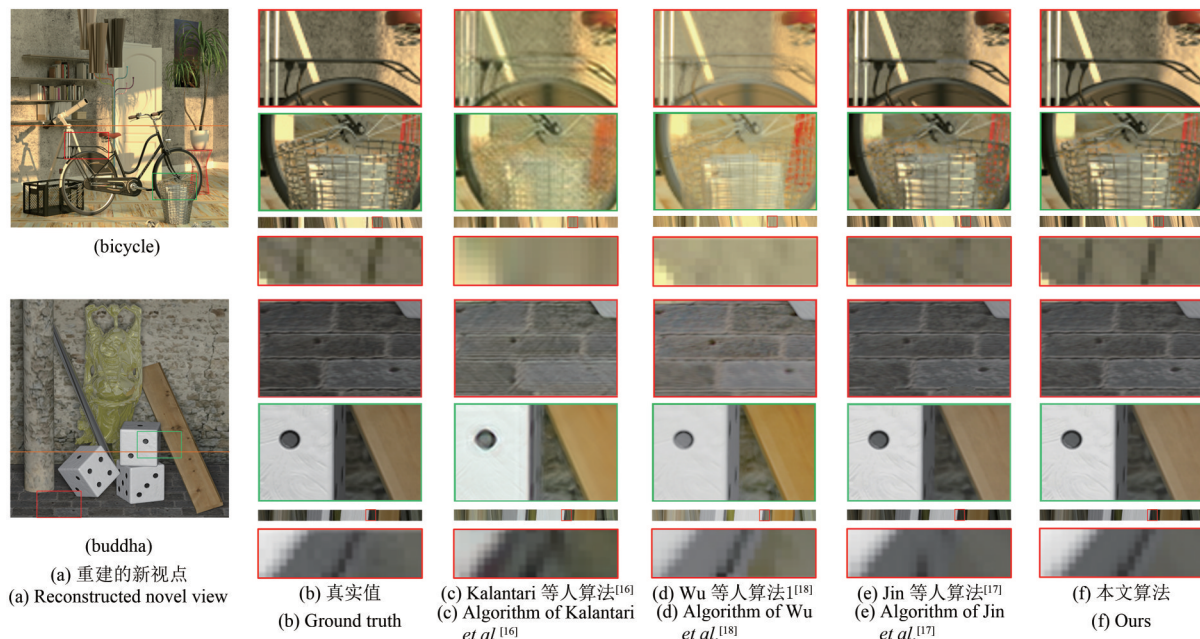


图 4 不同方法对合成光场数据集的重建结果

Fig. 4 Reconstruction results of synthetic light field data set by different methods

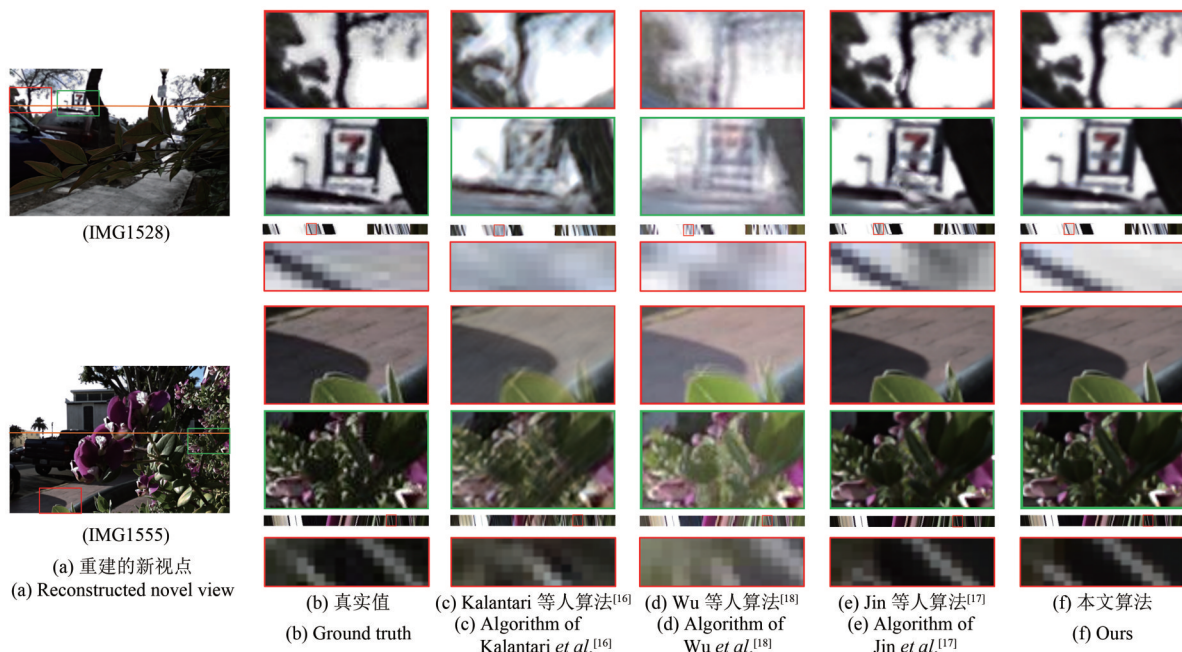


图 5 不同方法对真实世界数据集的重建结果

Fig. 5 Reconstruction results of real-world data sets by different methods

然显式地估计了深度图,但是他们的网络过于简单,仅使用4层顺序连接的卷积,同样在遮挡附近出现了模糊;Jin 等人^[17]使用了9层顺序连接的卷积估计深度图,在融合步骤使用了空间角度卷积来提取光场空间角度特征,在大多数场景取得了次优的效果。但是他们的深度估计模块未能利

用光场子孔径图像蕴含的多方向信息,不能很好地解决遮挡问题。从自行车中的后座边缘、IMG1528的树枝和 IMG1555 的树叶等场景可以看出,本文提出的方法能更好地应对各种遮挡问题。

从所有场景的EPI结果可以看出,本文提出算法的EPI在物体边缘等区域保留了更好的线型结

构,进一步验证了本文算法能较好地处理遮挡问题。

3.3 消融实验

本文通过多个消融实验来验证本文所提出的多路输入模式和空间角度解耦融合模块在光

场重建过程中所起到的作用和对最终结果的贡献程度,定性和定量地评估了本文的多路输入和空间角度解耦融合模块的有效性,给出了消融不同模块下重建光场的局部放大图。

表 4 不同模块的消融实验

Tab. 4 Ablation experiments of different modules

	HCI	HCI old	30 scenes
Single branch	34.60/0.937	40.84/0.960	41.47/0.983
Multi branches	34.96/0.941	41.02/0.960	41.65/0.984
Multi branches + SADFM	35.60/0.950	41.52/0.961	42.07/0.985

图 6 给出了重建中心子孔径图像局部放大图。可以看到多路输入使重建视图在遮挡区域保持良好的效果,但是空间纹理部分产生较粗糙的结果,并且会出现一些不适定区域。多路输入

加上解耦融合模块充分探索了光场空间角度内在关系,并以空间信息辅助角度重建,使结果在纹理重复区域更加清晰,并且消除了不适定区域。

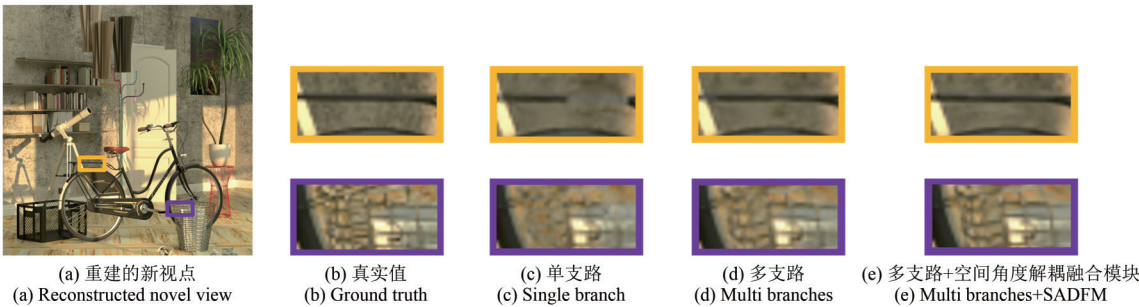


图 6 消融不同模块的对比效果

Fig. 6 Comparative effects of different ablation modules

4 结 论

本文提出了一种基于空间角度解耦融合的光场重建算法。本文方法充分利用光场蕴含的深度信息和神经网络强大的学习能力,通过显式估计场景深度和充分探索空间角度信息来实现

有效的大基线光场重建。实验结果表明,本文方法对比 Jin 等人^[17]的方法,在 HCI、HCI old 和 30 scenes 数据集上的平均 PSNR 分别提高了 1.0、0.68 和 0.6 dB,平均 SSIM 提高了 0.013、0.001 和 0.002。本文方法具有更高的重建质量,在遮挡区域、纹理重复区域保持了更好的重建效果。

参 考 文 献:

[1] 陈苑锋. 视觉深度估计与点云建图研究进展[J]. 液晶与显示, 2021, 36(6): 896-911.
CHEN Y F. Progress of visual depth estimation and point cloud mapping [J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(6): 896-911. (in Chinese)

[2] MENG N, LI K, LIU J Z, *et al.* Light field view synthesis via aperture disparity and warping confidence map [J]. *IEEE Transactions on Image Processing*, 2021, 30: 3908-3921.

[3] WU G C, WANG Y Q, LIU Y B, *et al.* Spatial-angular attention network for light field reconstruction [J]. *IEEE Transactions on Image Processing*, 2021, 30: 8999-9013.

- [4] KO K, KOH Y J, CHANG S, *et al.* Light field super-resolution via adaptive feature remixing [J]. *IEEE Transactions on Image Processing*, 2021, 30: 4114-4128.
- [5] CHEN A P, XU Z X, ZHAO F Q, *et al.* MVSNeRF: fast generalizable radiance field reconstruction from multi-view stereo [C]//*Proceedings of 2021 IEEE/CVF International Conference on Computer Vision*. Montreal: IEEE, 2021: 14104-14113.
- [6] SHI L X, HASSANIEH H, DAVIS A, *et al.* Light field reconstruction using sparsity in the continuous Fourier domain [J]. *ACM Transactions on Graphics*, 2014, 34(1): 12.
- [7] 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.
- [8] WU G C, ZHAO M D, WANG L Y, *et al.* Light field reconstruction using deep convolutional network on EPI [C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 1638-1646.
- [9] PENDU M L, GUILLEMOT C, SMOLIC A. A Fourier disparity layer representation for light fields [J]. *IEEE Transactions on Image Processing*, 2019, 28(11): 5740-5753.
- [10] 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 [C]//*Proceedings of the 15th European Conference on Computer Vision (ECCV)*. Munich: Springer, 2018: 138-154.
- [11] MENG N, SO H K H, SUN X, *et al.* High-dimensional dense residual convolutional neural network for light field reconstruction [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(3): 873-886.
- [12] MENG N, WU X F, LIU J Z, *et al.* High-order residual network for light field super-resolution [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7): 11757-11764.
- [13] JEON H G, PARK J, CHOE G, *et al.* Accurate depth map estimation from a lenslet light field camera [C]//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015: 1547-1555.
- [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] FLYNN J, NEULANDER I, PHILBIN J, *et al.* Deep stereo: learning to predict new views from the world's imagery [C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 5515-5524.
- [16] KALANTARI N K, WANG T C, RAMAMOORTHY R. Learning-based view synthesis for light field cameras [J]. *ACM Transactions on Graphics*, 2016, 35(6): 193.
- [17] 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.
- [18] WU G C, LIU Y B, DAI Q H, *et al.* Learning sheared EPI structure for light field reconstruction [J]. *IEEE Transactions on Image Processing*, 2019, 28(7): 3261-3273.
- [19] 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.
- [20] MILDENHALL B, SRINIVASAN P P, ORTIZ-CAYON R, *et al.* Local light field fusion: practical view synthesis with prescriptive sampling guidelines [J]. *ACM Transactions on Graphics*, 2019, 38(4): 29.
- [21] 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*. Taipei, China: Springer, 2017: 19-34.
- [22] WANNER S, MEISTER S, GOLDLUECKE B. Datasets and benchmarks for densely sampled 4D light fields [M]//BRONSTEIN M, FAVRE J, HORMANN K. *Vision, Modeling & Visualization*. Geneva: The Eurographics Association, 2013: 225-226.
- [23] KINGMA D P, BA J. Adam: a method for stochastic optimization [J]. *arXiv*, 2014: 1412.6980.

作者简介:



张洪基(1998—),男,湖北襄阳人,硕士研究生,2020年于武汉科技大学获得学士学位,主要从事光场图像处理、机器学习等方面的研究。E-mail: 1363209384@qq.com



邓慧萍(1983—),女,湖北应城人,博士,副教授,2013年于华中科技大学获得博士学位,主要从事3D视频与图像处理、机器学习等方面的研究。E-mail: denghuiping@wust.edu.cn