文章编号　1004-924X(2024)03-0456-10

联合傅里叶卷积与通道注意力的光场角度重建

周　涛[1]，郁　梅[1*]，陈晔曜[1]，蒋志迪[2]，蒋刚毅[1]
(1. 宁波大学 信息科学与工程学院，浙江 宁波 315211；
2. 宁波大学科学技术学院 信息工程学院，浙江 宁波 315212)

摘要：光场相机能够同时捕获光线的强度和方向信息，但由于成像传感器尺寸的限制，无法同时获得高空间和角度分辨率的光场图像。提出了一种联合傅里叶卷积和通道注意力的光场角度重建方法，通过使用稀疏光场图像4个边角位置的参考视图，可以间接地重建出密集光场图像。考虑到光场数据的内在4D结构，采用通道级密集快速傅里叶残差卷积块，在空域和频域对光场图像的空间和角度相关性进行建模，然后采用基于全局响应归一化的通道注意块，以实现通道间的自适应融合。此外，还提出了一种改进的视点加权间接合成方法，通过为每个参考视图分配一个置信图，为参考视图之间建立联系以合成更真实的新视图。实验结果表明，相比于现有先进的光场角度重建算法IRVAE，所提方法的重建光场图像质量在自然光场数据集30Scenes，Occlusion和Reflective上的平均PSNR分别提高了0.08，0.13和0.13 dB。所提方法在保证光场角度一致性的前提下取得了清晰的重建结果。

关　键　词：光场角度重建；傅里叶卷积；全局响应归一化；视点加权的间接合成
中图分类号：TP391.41　　**文献标识码**：A　　**doi**：10.37188/OPE.20243203.0456

Light field angular reconstruction with joint Fourier convolution and channel attention

ZHOU Tao[1], YU Mei[1*], CHEN Yeyao[1], JIANG Zhidi[2], JIANG Gangyi[1]

(1. *Faculty of Information Science and Engineering, Ningbo University, Ningbo* 315211, *China*;
2. *College of Information Engineering, College of Science and Technology Ningbo University, Ningbo* 315212, *China*)
* *Corresponding author, E-mail: yumei@nbu.edu.cn*

Abstract: Light field cameras, capable of capturing both the intensity and direction of light, are utilized in applications like foreground occlusion removal and depth estimation. However, the limited size of imaging sensors restricts the simultaneous achievement of high spatial and angular resolution in light field images. This paper introduces a method combining Fourier convolution and channel attention for light field angular reconstruction, which indirectly creates dense light field images from sparse ones using reference views from the image's four corners. This method leverages the light field image's inherent 4D structure, employing channel-level dense fast Fourier residual convolution blocks to model the spatial and angular correlations in both spatial and frequency domains. Channel attention blocks, utilizing global response normaliza-

收稿日期：2023-08-07；**修订日期**：2023-10-10.
基金项目：国家自然科学基金资助项目(No. 62071266，No. 62271276)；宁波市自然科学基金资助项目(No. 202003N4088)

tion, then adaptively fuse these channels. Furthermore, an enhanced viewpoint-weighted indirect synthesis approach is proposed, assigning a confidence map to each reference view to improve the synthesis of new, realistic views by establishing relationships between reference views. Experimental results demonstrate that our method outperforms the advanced light field angular reconstruction technique IRVAE, showing an average PSNR improvement of 0.08, 0.13, and 0.13 dB on the natural light field dataset 30Scenes, Occlusion, and Reflective, respectively, ensuring clear reconstruction results with angular consistency in the light field.

Key words: light field angular reconstruction; Fourier convolution; global response normalization; viewpoint weighting indirect view synthesis

1　引　言

区别于传统成像只能在单个方向上捕获三维空间的光线信息,光场成像技术能够同时记录场景中光线的强度和方向信息。基于光场成像的光学仪器(即光场相机)也被开发以获取更丰富的场景信息。许多光场应用也随之产生,如深度感知[1]、反射率估计[2]、视图渲染[3]、前景去遮挡[4]等技术。通过在主镜头和成像传感器之间插入微透镜阵列等光学组件,光场相机可以通过单次曝光同时采集空间信息和角度信息。但受限于传感器的尺寸,密集的空间采样会导致稀疏的角度采样,这严重阻碍了光场成像的实际应用。

为了解决这个问题,基于卷积神经网络(Convolutional Neural Network,CNN)的光场角度超分辨率算法被提出。但由于光场图像的四维(4-Dimensions,4D)结构限制,其空间信息与角度信息高度耦合,给卷积神经网络的光场应用带来了挑战。现有的基于卷积神经网络的方法通过直接生成或者间接生成两种方式来获得密集的光场图像。

直接生成法先从稀疏光场图像中建模空间和角度信息的相关性,再沿角度维上采样重建光场。Yoon等[5]首次用CNNs对光场图像建模,通过邻域视图建模的方法从相邻的两个子孔径图(Sub-Aperture Image, SAI)中生成中间视图。Yeung等[6]提出空间角度可分离卷积来代替4D卷积提取光场4D结构信息。Wu等[7]将极平面图像(Epipolar Plane Image,EPI)视为光场图像的基本单元,提出基于EPI的重建网络,但因EPI本身分辨率的问题,该网络在低角度分辨率作为输入的情况表现欠佳。Wang等[8]提出一个端到端的伪4D CNN,将二维(2-Dimensions,2D)EPIs堆叠成三维(3-Dimensions,3D)形式作为输入进行角度重建。Wang等[9]将光场图像视为宏像素图像阵列,并设计了一种解耦机制来充分利用光场的角度信息。间接生成法大多通过生成一些中间输出,通过中间输出与输入的操作来重建光场图像。Kalantari等[10]提出一个端到端的两阶段网络,将角度重建看作视差估计和色彩估计两部分,在生成中间输出视差图后,根据输入与视差图绘制出粗糙结果,后续进行色彩补偿。Wu等[11]通过预移位的EPIs隐式地估计场景深度,并提出一种克服EPI不匹配的CNN重建网络,可实现更大视差范围下的光场重建。除此之外,Jin等[12]提出一个能从非结构化稀疏光场输入重建出密集分布的两阶段网络。上述直接和间接方法都只能生成密集分布的光场图像,无法从稀疏分布的光场图像中重建出任意角度位置的新视图。近期,Han[13]等提出一个基于变分自编码器的间接生成网络,它能够从稀疏分布的光场输入图像中为每个参考视图生成一组非共享卷积核,通过与参考视图的卷积可以灵活地得到任意角度位置的新视图。但它与其他角度超分方法存在一样的问题,即特征提取时受限于感受野,在更大尺寸光场图像上对空间和角度信息的相关性建模不充分。

为了解决上述问题,本文提出了一个简单有效的方法来调整光场空角相关性建模时的感受野。鉴于频域上的一点能影响空域上的全局信息、频域的全局信息与空间上局部信息存在相关性,基于快速傅里叶卷积[14]提出了一个密集快速傅里叶卷积残差(Dence Fast Fourier Convolutions Residual,DFFCR)块来更有效地建模光场

的空间和角度相关性。该模块分别在频域和空域上进行了卷积操作，以提取场景的全局和局部信息。同时，通过引入基于全局响应归一化(Global Response Normalization，GRN)[15]的通道注意块，能够将全局信息与局部信息进行通道级融合，更有效地利用光场图像的空间和角度信息。其次，提出了一种视点加权的间接合成(Viewpoint Weighting Indirect View Synthesis，VWIVS)块，该块能结合多个参考视图以生成最终结果。为每个参考视图生成置信图，并根据置信图来决定每个参考视图生成结果的权重。将每个参考视图生成结果进行融合后，得到最终输出。这一策略能够保留更多的细节信息，增强生成结果的可视化效果。

2　原　理

基于双平面光场参数化模型[16]，光场图像通常表示为一个 4D 函数 $L(u,v,s,t)\in \boldsymbol{R}^{U\times V\times S\times T}$，其中 U 和 V 表示角度维度，S 和 T 表示空间维度，在角度位置 (u,v) 上的 SAI 表示为 $I_{(u,v)}(s,t)\in\mathbf{R}^{S\times T}$，与自然 2D 图像具有相似的风格。

本文旨在从稀疏分布的参考子孔径图重建出新角度位置上的 SAI，使其尽可能接近真值。即给定输入参考子孔径图 L_{ref} 和目标角度位置 p_{tar}，该问题可以表示为：

$$\hat{I}_{(u_{\mathrm{tar}},v_{\mathrm{tar}})}(s,t)=f(L_{\mathrm{ref}},p_{\mathrm{tar}}),\tag{1}$$

其中：$\hat{I}_{(u_{\mathrm{tar}},v_{\mathrm{tar}})}(s,t)\in\mathbf{R}^{S\times T}$ 表示角度位置 $(u_{\mathrm{tar}},v_{\mathrm{tar}})$ 上的 SAI；$L_{\mathrm{ref}}=\{I_{(u',v')}(s,t)\}\in\mathbf{R}^{2\times 2\times S\times T}$，表示边角位置的 SAIs，如图 1 所示；$f(\cdot)$ 表示所提方法。

图 1 为所提方法的整体框架。重建过程主要包括初始特征提取模块、空频域特征学习模块、目标角度位置特征映射模块和视点加权的间接视图合成模块 4 个模块。首先，利用初始特征提取模块结合通道注意块初步提取参考子孔径图的空间信息。之后结合空域和频域上的卷积对参考子孔径图的空间和角度信息进行融合。结合目标角度位置后，将融合后的特征映射至目标角度位置，利用带有目标角度信息的特征通过映射模块为每个参考子孔径图的每个像素生成非共享卷积核，最后用该卷积核和参考子孔径图间接合成高质量且细节丰富的目标角度位置子孔径图。

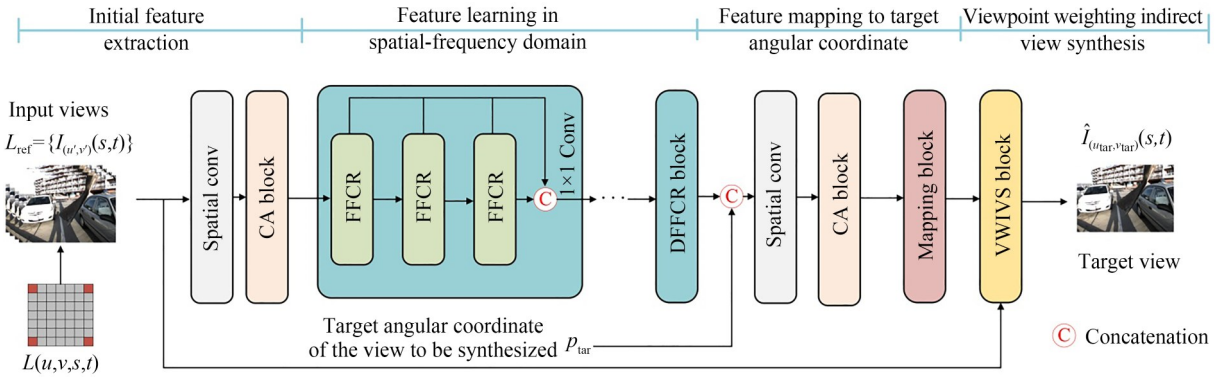


图 1　联合傅里叶卷积与通道注意力的光场重建方法的总体框图

Fig. 1　Framework of light field reconstruction method with joint Fourier convolution and channel attention

2.1　初始特征提取

首先，使用由少量 3×3 卷积加上激活层构成的 Spatial Conv 块将参考子孔径图映射至特征维度，如图 2(a)所示。为了在空间维度上更好地融合参考子孔径图之间的信息，结合基于 GRN 的通道注意块进一步融合参考子孔径图间的信息，以便在不产生额外参数的情况下增加通道间的对比和选择性，如图 2(b)所示。其中，使用 K 个级联的 ConvNeXt v2[15]块来实现对参考子孔径图在特征域中的信息融合。初步提取的特征表示为 $F\in\mathbf{R}^{C\times S\times T}$，其中 C 表示通道维度。

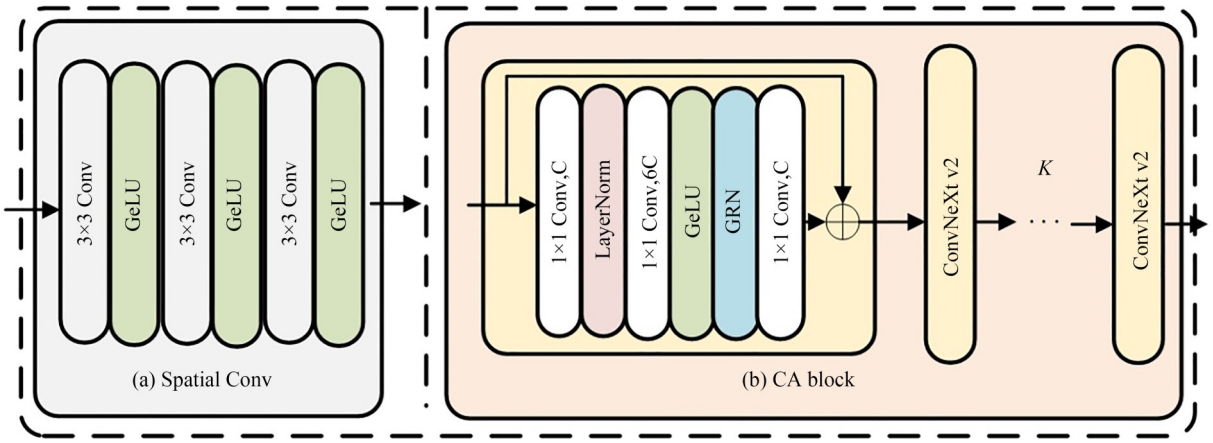


图 2　初始特征提取模块示意图

Fig. 2　Schematic diagram of initial feature extraction module

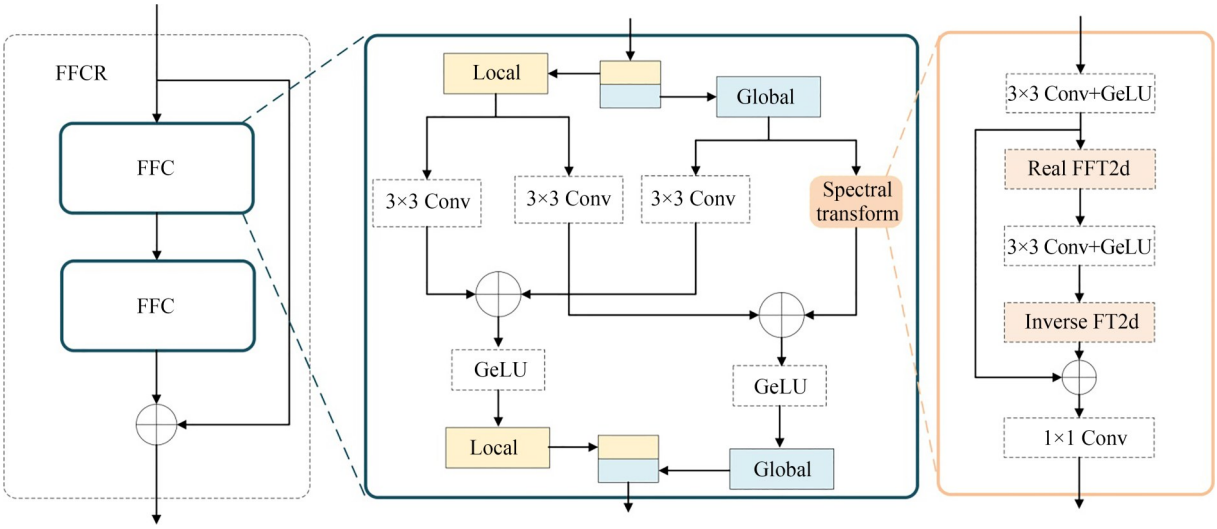


图 3　快速傅里叶卷积残差块示意图

Fig. 3　Schematic diagram of FFCR block

2.2　空频域特征学习

为整合多级特征学习与傅里叶卷积,设计了DFFCR块以提取子孔径图间的空域和频域信息。如图1所示,每个DFFCR块由3个级联的快速傅里叶卷积残差(Fast Fourier Convolutions Residual,FFCR)块和一个1×1卷积块组成,前两个FFCR块的输出会拼接至最后一个FFCR块,并通过1×1卷积块进行融合。假定$F_l^s \in \mathbf{R}^{C\times S\times T}$表示第$s$个DFFCR块内的第$l$个FFCR的输出,那么第$s$个DFFCR块的输出可以表示为:

$$F_{\text{out}}^s = W_c^s([F_1^s, F_2^s, F_3^s]), \tag{2}$$

其中W_c^s表示第s个DFFCR块内的1×1卷积块权重。适量的多级特征学习保证了各级前向特征的有效融合,也缓解了因为空域和频域信息的差异导致融合时不稳定的问题。

如图3所示,每个FFCR块包含两个快速傅里叶卷积(Fast Fourier Convolution,FFC)块。FFC块是基于通道级的快速傅里叶变换,它将输入特征沿着通道维度划分为局部和全局两个部分分别进行处理。局部分支使用普通的卷积来捕获局部特征;全局分支则利用一个频域变换块,在频域上考虑图像的全局结构并提取非局部信息。最终两个分支的输出堆叠在一起进行输出。频域变换块使用傅里叶卷积单元来提取全局信息。傅里叶卷积单元中主要使用Real FFT2d将输入从空域变换至频域中,然后在频域上进行卷积操作,最后使用Inverse FFT2d将特征恢复至空域。

2.3 目标角度位置特征映射

经过空频域特征学习后的输出特征只是对输入的参考子孔径图的空间和方向信息建模，还需要将它映射至角度位置。因此，对于给定目标角度位置P_{tar}，使用一个空间卷积块卷积W_{sc}进行初步融合。融合角度过程可以表示为：

$$F_{\mathrm{fused}}=W_{\mathrm{sc}}([\,p_{\mathrm{tar}},F_{\mathrm{DFFCR}}\,]),\tag{3}$$

其中，$F_{\mathrm{fused}}\in\mathbf{R}^{C\times S\times T}$表示初步融合角度后的输出，$F_{\mathrm{DFFCR}}\in\mathbf{R}^{C\times S\times T}$表示DFFCR输出的特征。由于DFFCR块和角度融合都是通道级别的，需要解决如何在模型稳定的情况下，有效地融合目标角度位置和所提取特征的问题。为此，采用一个与初始特征提取过程相同结构但不共享权重的通道注意力(Channel Attention，CA)块，稳定地融合提取特征和目标角度位置。参考现有的光场灵活角度位置重建工作[14]，使用残差密度块(Residual Dense Block，RDB)[17]将输入映射至目标卷积核。

2.4 视点加权的间接视图合成

现有的光场间接视图合成方法[13]先用自适应卷积[18]得到参考子孔径图的合成结果，再用相加的方式得到最终的子孔径图。这种融合方式不能保留真实的细节。本文借鉴立体匹配研究[19]，在最终融合过程中加入一个中间操作，通过置信图的方式调整参考子孔径图合成结果间的关系，以获得更真实的图像。

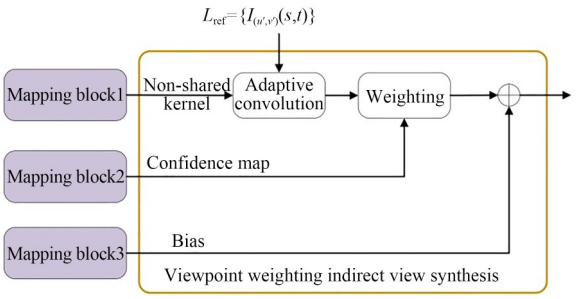


图4 视点加权的间接视图合成示意图

Fig. 4 Schematic of viewpoint weighting indirect view synthesis module

图4为每个参考子孔径图自适应卷积融合的结果分配一个像素级的置信图，在最终融合的过程中通过加权每个参考子孔径图的结果，辅以全局残差得到最终的目标子孔径图。考虑到l_1损失函数对异常值稳定，采用l_1损失函数来最小化重建子孔径图与真值(Ground Truth，GT)之间的平均绝对误差：

$$l=\frac{1}{n}\sum_{p=1}^{n}\left|I_{\mathrm{gt}}-\hat{I}\right|,\tag{4}$$

其中：n表示图像的像素总数，I_{gt}代表GT子孔径图，$\hat{I}$表示网络重建的子孔径图。

3 实验结果与分析

3.1 数据集及实现细节

基于文献[9]中的策略，使用自然光场数据集30Scenes[10]，STFlytro[20]进行实验。自然光场图像通常具有较小的基线，即相邻子孔径图视差较小，所有场景的光场图像的角度分辨率为14×14，空间分辨率为376×541。由于光场相机成像特性，光场图像的边缘子孔径图通常并不完整，因此，所有的光场图像都取其中心7×7的子孔径图作为参考的高角度分辨率光场图像。对于每个光场图像，选择2×2的角子孔径图作为输入的低角度分辨率光场图像。训练集和测试集划分如表1所示，使用30Scenes数据集中的100个自然光场图像用作训练。测试集则由30个选自30Scenes数据集的光场图像以及STFlytro数据集的15个Reflective场景和25个Occlusion场景的光场图像构成，训练集和测试集互不相交。在训练过程中，每个子孔径图被裁剪成64×64的图像块。测试则使用完整子孔径图。

采用YCbCr颜色空间中Y通道的峰值信噪比(Peak Signal-to-Noise Ratio，PSNR)和结构相似度(Structure Similarity Index Measure，SSIM)来衡量合成结果的客观质量。由于所提出的方法能够合成任意角度位置的新视图，首先计算所有位置的合成结果(即由光场图像2×2边角位置的SAIs生成的7×7共45个新视图)的PSNR和SSIM，然后取其平均值作为该光场图像的客观结果。此外，数据集的PSNR和SSIM是所有光场图像结果的平均值。

所有实验基于Pytorch深度学习框架完成，实验环境配置为24 vCPU Intel(R) Xeon(R) Platinum 8255C CPU @ 2.50GHz，两张RTX 3090(24GB)显卡。采用Adam算法作为优化器，初始学习率设置为0.000 2，并采用周期为60ep-

表 1　实验所用训练和测试集划分

Tab. 1　Partition of training and testing sets in experiments

Using	Dataset	Data type	Number of scenes
Training	30Scenes	Real-world	100
	Reflective	Real-world	15
Testing	Occlusion	Real-world	25
	30Scenes	Real-world	30

och的余弦退火优化策略。训练和测试过程与文献[14]一致。

3.2　与现有方法在基准数据集上的比较

为验证所提方法的有效性,采用ShearedEPI[11],Yeung[6],LFASR-geo[21],FS-GAF[12],DistgASR[9]和IRVAE[13]等进行对比实验。其中,IRVAE和所提方法均为灵活角度位置的重建方法。公平起见,所有方法都是在相同的数据集上进行训练。表2给出了对比实验结果,其中最好的性能指标用粗体标记。

由于稀疏光场图像的EPI仅包含2个像素行或像素列,很难重建光场图像中间的线性结构,因此如表2所示,基于EPI的方法[11]性能不如其他方法。相比之下,基于深度估计的方法如LFASR-geo[21]和FS-GAF[12]取得了优于基于EPI方法的性能。DistgASR[9]通过将光场结构解耦成4个2D分支进行多维信息融合直接重建缺失的视图,在真实场景上取得了比基于视差估计方法更好的性能。IRVAE[13]通过变分自编码器生成非共享卷积核间接合成任意缺失视图,取得比前两类方法更好的性能。所提出的方法通过结合光场的空频域信息学习光场的空间角度相关性以重建缺失的视图,在真实场景的所有数据集上取得了最好的性能指标。

图5展示了不同方法重建的缺失视图的主观视觉效果,重建视图在光场图像中的角度位置如图5(a)所示。图5(a)表示30Scenes数据集中的IMG_1554(上)和IMG_1541(下)光场图像重建视图对应角度位置的真值视图,图5(b)~5(e)给出了用不同方法重建的视图相对于真实视图的误差图,同时也给出了对应的2处局部放大结果以及一幅EPI图像。从误差图可以看出,所提方法相比其他方法更接近真值,能够很好地还原场景的细节结构,如IMG_1541场景中草尖的轮廓形状。如图5局部放大图所示,该方法可以较好地从参考视图恢复出目标视图的颜色以及纹理细节,而对比方法在这些细节处产生失真。

为了展示密集分布光场图像的重建方法与灵活位置光场图像的重建方法的差异,图6进一步展示了DistgASR与所提方法在数据集30 Scenes上重建的各个SAIs的PSNR分布图。DistgASR为当前性能最好的密集分布光场图像的重建方法,方格中的数字代表对应角度位置上所有场景的光场SAI重建结果与其GT之间的平均PSNR。由图可以看出,DistgASR与所提方法在距离参考视图近的角度位置重建性能较好;而距离参考视图较远的位置如中心SAI,两种方法的重建性能相对略差,但也在42.7 dB之上。所提方法的重建性能在除少数距离参考试图较近的位置外均优于DistgASR,这说明它能更好地建模光场图像的空间和角度相关性。

表 2　不同光场角度重建方法在2×2→7×7任务上的PSNR和SSIM值

Tab. 2　PSNR and SSIM of different light field angular reconstruction methods on task of 2×2→7×7

Method	30 Scenes		Occlusion		Reflective	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
ShearedEPI[11]	39.17	0.975 4	34.41	0.954 8	36.38	0.944 3
Yeung[6]	42.77	0.985 6	38.88	0.979 7	38.33	0.960 2
LFASR-geo[21]	42.53	0.984 8	38.36	0.976 6	38.20	0.955 3
FS-GAF[12]	42.75	0.986 1	38.51	0.978 8	38.35	0.957 2
DistgASR[9]	43.64	0.995 1	39.46	0.990 8	39.11	0.978 0
IRVAE[13]	43.70	0.995 1	40.45	0.992 5	39.81	**0.982 3**
Proposed method	**43.78**	**0.995 2**	**40.58**	**0.992 7**	**39.94**	0.982 2

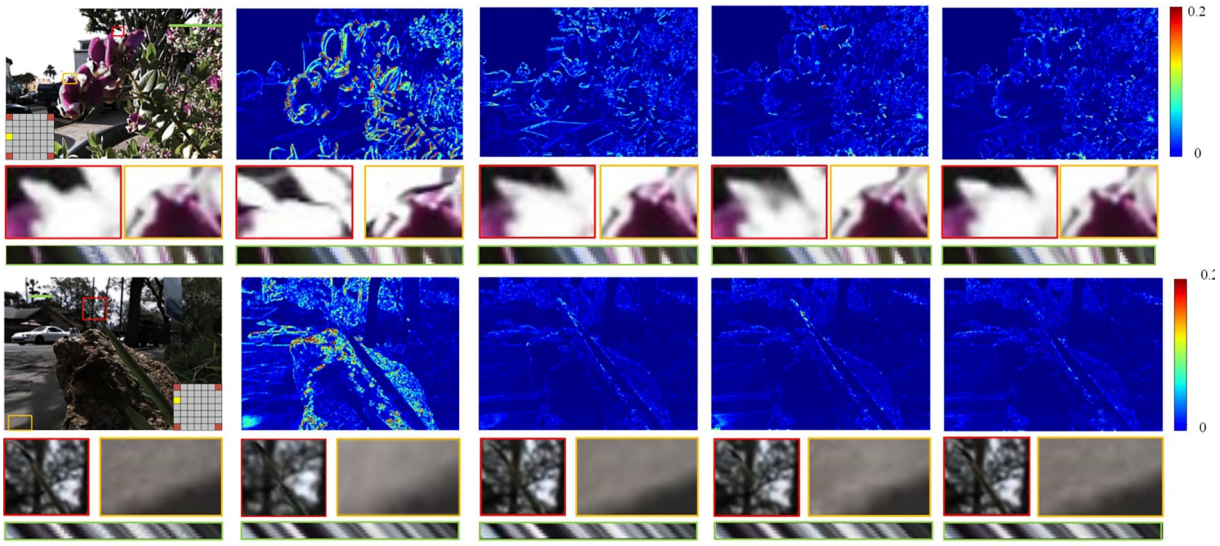


图 5　不同方法重建的缺失视图的主观视觉效果

Fig. 5　Subjective visual effects of missing views reconstructed by different methods

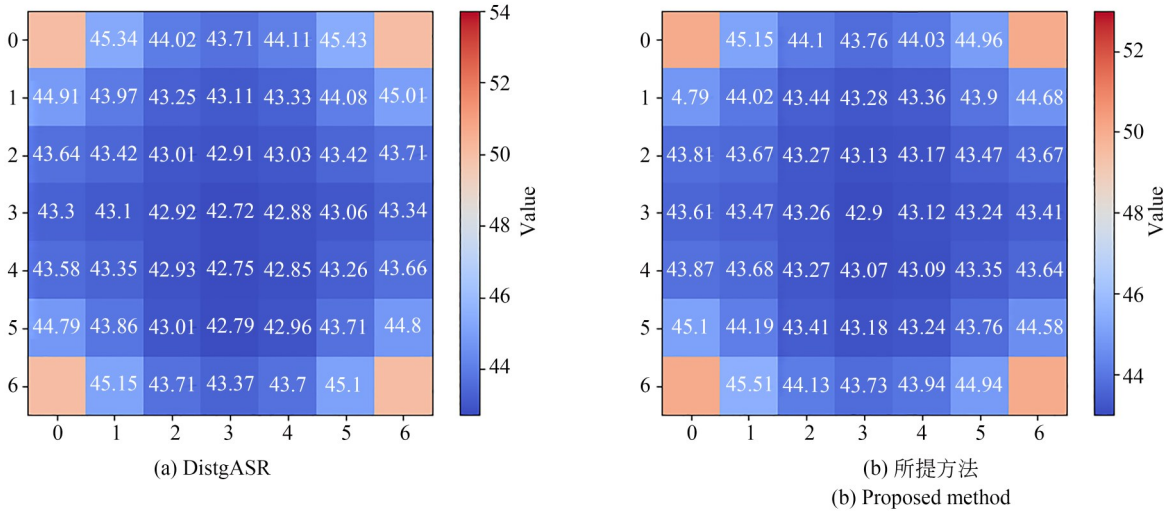


图 6　2×2→7×7 任务上 DistgASR [9] 和所提方法在数据集 30 Scenes 上重建的 SAIs 的 PSNR 分布

Fig. 6　PSNR distribution of SAIs achieved by DistgASR[9] and proposed method on 30 Scenes dataset on task of 2×2→7×7

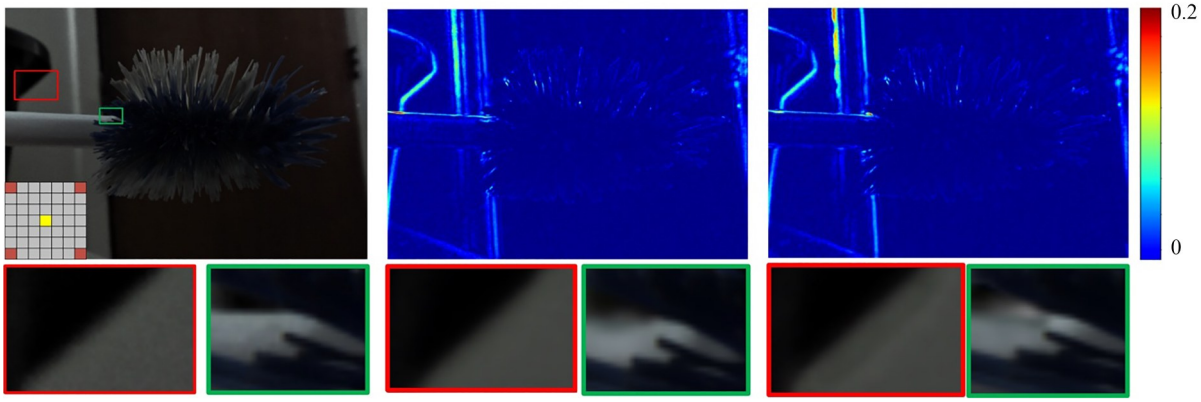
3.3　消融实验

选择性地从所提方法中删除 DFFCR，CAB 和 VWIVS 块，以验证各个块的有效性。表 3 为消融实验结果。如表 3 所示，对于前二者而言，缺少其中任意一个均会造成模型性能的下降，这归因于 DFFCR 块是通道级的，缺少 CAB 块的通道级特征融合会导致光场图像特征利用不充分；也证明 DFFCR 块融合空频域特征的有效性。其次，缺少 VWIVS 块会导致模型在所有数据集上的性能略微下降，说明联合参考视图进行融合会带来更好的结果。此外，通过对比所提方法是否包含 DFFCR 块来验证空频域信息充分结合对光场空间和角度信息建模的有效性。图 7 给出了所提方法中空频域特征学习模块对重建视图的影响。这里展示了重建出的中心子孔径图的误差图以及两个局部放大图。由图可知，带有 DFF-

表 3　所提方法在 2×2→7×7 任务上的消融实验

Tab. 3　Ablation experiments of proposed method on task of 2×2→7×7

DFFCR	CAB	VWIVS	30 Scenes		Occlusion		Reflective	
			PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
	√	√	43.56	0.994 8	40.10	0.991 3	39.79	0.981 8
√		√	43.28	0.994 6	39.80	0.991 2	39.63	0.980 4
√	√		43.66	0.995 1	40.21	0.991 6	39.81	0.981 4
√	√	√	**43.78**	**0.995 2**	**40.58**	**0.992 7**	**39.94**	**0.982 2**



(a) 原始中心子孔径图
(a) Original central subaperture image

(b) 所提方法的重建结果
(b) Reconstraction result of proposed method

(c) 不带DFFCR块方法的重建结果
(c) Reconstraction result of proposed method without DFFCR block

图 7　所提方法的 DFFCR 块在光场图像 IMG_1743 上的有效性视觉验证

Fig. 7　Visual verification of validity of DFFCR block in proposed method on light field image IMG_1743

CR 块的方法相比不带 DFFCR 块的方法的误差更小。

4　结　论

本文提出了一种联合傅里叶卷积和通道注意力的间接视图合成方法,通过合成任意角度位置的新视图间接进行光场角度重建。该方法包括初始特征提取、空频域特征学习、目标角度位置特征映射和视点加权的间接视图合成,获得了比一些先进方法更真实的结果和富有高频信息的结构。实验结果表明,相比 IRVAE,所提方法的重建光场图像质量在自然光场数据集 30Scenes,Occlusion 和 Reflective 上的平均 PSNR 分别提升了 0.08,0.13 和 0.13 dB,综合性能优于现有方法。所提出的方法在保证光场角度一致性的前提下取得了清晰的重建结果。但本文只能从固定分布的参考子孔径图重建任意角度位置的新视图,在面向灵活输入分布、灵活输入数量重建问题时无法以单模型应对。在未来的工作中,将研究有效结合空频域信息对光场图像进行更合理建模的方法,以及面向光场可伸缩编码的更灵活的光场重建方法。

参考文献:

[1] WANG Y, WANG L, LIANG Z, *et al*. Occlusion-aware cost constructor for light field depth estimation [C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*). 18-24, 2022, *New Orleans*, *LA*, *USA*. IEEE, 2022: 19777-19786.

[2] ZHOU M Y, DING Y Q, 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.

[3] ATTAL B, HUANG J B, ZOLLHÖFER M, *et al*. Learning neural light fields with ray-space embedding[C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*). 18-24, 2022, *New Orleans*, *LA*, *USA*. IEEE, 2022: 19787-19797.

[4] LI Y J, YANG W, XU Z B, *et al*. Mask4D: 4D convolution network for light field occlusion removal [C]. *ICASSP* 2021 - 2021 *IEEE International Conference on Acoustics*, *Speech and Signal Processing* (*ICASSP*). *June* 6-11, 2021. *Toronto*, *ON*, *Canada*. IEEE, 2021: 2480-2484.

[5] 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* (*ICCVW*). ACM, 2015: 57 - 65.

[6] 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]. *Computer Vision-ECCV* 2018: 15*th European Conference*, *Munich*, *Germany*, *September* 8-14, 2018, *Proceedings*, *Part VI*. ACM, 2018: 138 - 154.

[7] WU G C, ZHAO M D, WANG L Y, *et al*. Light field reconstruction using deep convolutional network on EPI[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*). *July* 21-26, 2017. *Honolulu*, *HI*. IEEE, 2017: 6319-6327.

[8] WANG Y L, LIU F, WANG Z L, *et al*. *End-to-end View Synthesis for Light Field Imaging with Pseudo 4DCNN* [M]. Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 340-355.

[9] WANG Y, WANG L, WU G, *et al*. Disentangling light fields for super-resolution and disparity estimation[J]. *IEEE Trans Pattern Anal Mach Intell*, 2023, 45(1): 425-443.

[10] KALANTARI N K, WANG T C, RAMAMOORTHI R. Learning-based view synthesis for light field cameras [J]. *ACM Transactions on Graphics*, 35(6): 193.

[11] 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.

[12] JIN J, HOU J H, CHEN J, *et al*. Deep coarse-to-fine dense light field reconstruction with flexible sampling and geometry-aware fusion [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(4): 1819-1836.

[13] HAN K, XIANG W. Inference-reconstruction variational autoencoder for light field image reconstruction[J]. *IEEE Transactions on Image Processing*, 2022, 31: 5629-5644.

[14] SUVOROV R, LOGACHEVA E, MASHIKHIN A, *et al*. Resolution-robust large mask inpainting with Fourier convolutions [C]. 2022 *IEEE/CVF Winter Conference on Applications of Computer Vision* (*WACV*). *January* 3-8, 2022. *Waikoloa*, *HI*, *USA*. IEEE, 2022: 2149-2159.

[15] WOO S, DEBNATH S, HU R H, *et al*. ConvNeXt V2: co-designing and scaling ConvNets with masked autoencoders[C]. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*). *June* 17-24, 2023. *Vancouver*, *BC*, *Canada*. IEEE, 2023: 16133-16142.

[16] LEVOY M, HANRAHAN P. Light field rendering[C]. *Proceedings of the* 23*rd annual conference on Computer graphics and interactive techniques*. ACM, 1996: 31 - 42.

[17] ZHANG Y L, TIAN Y P, KONG Y, *et al*. Residual dense network for image super-resolution [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. *June* 18-23, 2018. *Salt Lake City*, *UT*. IEEE, 2018: 2472-2481.

[18] NIKLAUS S, MAI L, LIU F. Video frame interpolation via adaptive convolution[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*). *July* 21-26, 2017. *Honolulu*, *HI*. IEEE, 2017: 670-679.

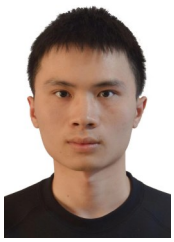
[19] CHEN L Y, WANG W H, MORDOHAI P. Learning the distribution of errors in stereo matching for joint disparity and uncertainty estimation [C]. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*). *June* 17-24, 2023. *Vancouver*, *BC*, *Canada*. IEEE, 2023: 17235-17244.

[20] RAJA S, LOWNEY M, SHAH R, *et al*. Stanford lytro light field archive[J]. *LF*2016. *html*, 2016.

[21] 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.

作者简介：



周　涛(2000—)，男，湖南衡阳人，硕士研究生，2021年于南昌大学获得学士学位，主要从事光场超分辨率重建等方面的研究。E-mail：794543231@qq.com

通讯作者：

郁　梅(1968—)，女，江苏无锡人，博士，教授，博士生导师，2000年于韩国Ajou大学获得博士学位，主要从事多媒体信号处理与通信、计算成像、视觉感知与编码、图像与视频质量评价等方面的研究。E-mail：yumei@nbu.edu.cn