

文章编号 1004-924X(2023)21-3167-11

## 融合 3D 对极平面图像的光场角度超分辨重建

陈纯毅\*, 范晓辉, 胡小娟, 于海洋

(长春理工大学 计算机科学技术学院, 吉林 长春 130022)

**摘要:**针对光场成像中因硬件限制而造成的光场图像角度分辨率低的问题,提出一种融合 3D 对极平面图像的光场角度超分辨重建方法。该方法首先将输入图像按不同的视差方向排列分别进行特征提取,以充分利用输入图像的视差信息,提高深度估计的准确性。利用深度图将输入图像映射到新视角位置,生成初始合成光场。为了使重建光场图像能够保持更好的细节信息及几何一致性,先通过水平 3D 对极平面图像融合重建分支和垂直 3D 对极平面图像融合重建分支,分别对初始合成光场进行水平融合重建和垂直融合重建,再将两个结果进行混合重建,生成最终的高角度分辨率光场图像。实验结果表明:相比于现有方法,本文方法在合成光场数据集和真实光场数据集上的重建效果均得到了提高,峰值信噪比的提升幅度最高达 1.99%,有效地提高了重建光场的质量。

**关键词:**光场;超分辨重建;3D 对极平面图像;卷积神经网络

**中图分类号:**O439;TP391.41 **文献标识码:**A **doi:**10.37188/OPE.20233121.3167

## Light-field angular super-resolution reconstruction via fusing 3D epipolar plane images

CHEN Chunyi\*, FAN Xiaohui, HU Xiaojuan, YU Haiyang

(School of Computer Science and Technology, Changchun University of Science and Technology,  
Changchun 130022, China)

\* Corresponding author, E-mail: chenchunyi@hotmail.com

**Abstract:** Light-field (LF) imaging can capture spatial and angular information of light rays in a scene. Compared to traditional 2D/3D images, LF images provide a more comprehensive description of the scene. To address the problem of low angular resolution in LF images caused by hardware constraints, a LF angular super-resolution reconstruction method based on the fusion of 3D epipolar plane images (EPIs) is proposed. First, to make full use of the parallax information of the input images and improve the accuracy of depth estimation, the input images are arranged in varying parallax directions, and their features are extracted individually. Then, initial synthetic LF images are generated by transforming the input images to match the new viewpoint location using the corresponding depth maps. Finally, to ensure that the reconstructed LF image retains better detail information and geometric consistency, the LF is reconstructed horizontally and vertically via the horizontal and vertical 3D EPI fusion reconstruction branches, respectively. These two reconstruction results are then fused to produce the final high-angular-resolution LF image. Experimental results show that, compared to existing methods, the proposed method achieves an improved

收稿日期:2023-01-03;修订日期:2023-02-11.

基金项目:国家自然科学基金资助项目(No. U19A2063)

reconstruction quality across both synthetic and real-world LF image datasets, and the maximum increase in the peak signal-to-noise ratio reaches 1.99%. Thus, the proposed method can effectively improve the quality of the reconstructed LF.

**Key words:** light field; super-resolution reconstruction; 3D epipolar plane image; convolution neural network

## 1 引言

光场(Light Field, LF)成像可以捕获到现实世界的四维信息<sup>[1]</sup>(二维空间信息和二维角度信息),克服了传统成像技术只能捕获到二维空间信息的局限性,具有能在不同焦距下进行重投影的独特优势,逐渐在工业探测、生命科学和虚拟现实等领域受到重视<sup>[2-4]</sup>。光场图像的角度分辨率会直接影响三维重建、光场渲染等视觉应用的效果。然而,微透镜阵列光场相机受到传感器分辨率的限制,必须在空间分辨率和角度分辨率之间进行平衡。所以,提高光场图像的空间分辨率和角度分辨率一直是光场成像的研究热点。

光场角度超分辨重建(Angular Super-resolution Reconstruction, ASR)分为不基于深度信息的重建和基于深度信息的重建两种方法。不基于深度信息的重建方法在没有场景几何信息的情况下直接对角度维度进行上采样。Shi等<sup>[5]</sup>通过优化连续傅里叶域中的稀疏性进行光场重建。Vagharshakyan等<sup>[6]</sup>提出了一种自适应迭代正则化算法,在剪切波域中利用对极平面图像(Epipolar Plane Image, EPI)的稀疏表示来重建光场图像。但是,这些传统方法需要大量的输入图像。Yoon等<sup>[7]</sup>采用深度学习方法重建高分辨率光场,利用相邻的两个进行空间上采样后的图像逐个重建中间视角图像。该方法仅利用较少的角度信息,无法直接在任意位置合成新视图。Wu等<sup>[8]</sup>将光场重建问题转化为EPI的角度域细节恢复问题,提出“模糊-恢复-去模糊”的框架。该方法有效地提高了遮挡区域及非朗伯表面的重建效果,但是需要重复执行很多次。Wang等<sup>[9-10]</sup>构建了伪4D卷积神经网络(Convolutional Neural Network, CNN),在3D EPI上直接进行插值重建。Salem等<sup>[11]</sup>将角度维度上采样后的图像堆栈输入残差通道注意力网络,学习重要特征,恢复更多的纹理细节信息。由于没有建模场景的几何信息,这些方法都不能很好地处理更加稀疏采

样的输入图像。

基于深度信息的重建方法是指在重建过程中估计出所有视角的深度图,然后利用深度图将输入图像映射到新视角位置。Wanner等<sup>[12]</sup>利用结构张量法估计视差图,并通过能量函数惩罚映射图像和真实图像之间的误差来合成新视图,但这种方法必须在视点采样足够密集时才能达到合理的精度。Penner等<sup>[13]</sup>在软3D重建过程中充分考虑了深度信息的不确定性,使这个框架同时适用于结构化和非结构化的输入图像。传统方法一般聚焦于深度图的精度。近年来,基于深度学习的方法得到了广泛的应用。Kalantari等<sup>[14]</sup>预定义一组深度等级,并将每个等级上所有映射图像的均值和标准差作为手动提取到的特征输入端到端的顺序网络,进行深度估计和颜色估计。该方法有效提高了重建光场图像的质量,但是每次只能合成一个新视角图像。Vadathya等<sup>[15]</sup>利用散焦图像和聚焦图像进行视差估计,并通过修复网络修复遮挡造成的错误像素。Jin等<sup>[16]</sup>提出在空间域和角度域执行交错卷积的策略,有效提高了大基线光场的重建质量。Navarro等<sup>[17]</sup>在每个新视角位置对映射图像计算融合权重,并将其加权融合。Gul等<sup>[18]</sup>在融合映射图像时引入注意力机制,有效地处理了遮挡区域。Yun等<sup>[19]</sup>提出基于多重感受野的分层特征提取方法,用于提高深度估计的准确度。上述方法在处理映射图像时或只利用光场的空间信息,或交错利用空间信息和角度信息,不利于提高光场图像的重建质量。

光场图像的重建质量与能否充分利用光场图像的丰富信息有关。光场图像的3D EPI既包含空间信息,也包含角度信息<sup>[20]</sup>,利用这一特点,本文提出融合3D EPI的光场角度超分辨重建方法。该方法由3部分组成:(1)深度估计模块将输入图像按视差方向分为不同的图像对,分别进行特征提取,提取到的特征用来无监督地估计密集光场所有视角的深度图;(2)图像映射部分利用

得到的深度图将输入图像反向映射到每一个新视角位置,得到初始合成光场;(3)光场融合重建模块分为两个分支,分别从水平方向和垂直方向对初始合成光场的 3D EPI 进行融合,得到水平方向和垂直方向的重建结果,再对两个重建结果进行混合重建,得到最终的高角度分辨率光场图像。该方法通过评估 3D EPI 来引导光场融合,能够在处理映射图像时同时考虑空间信息和角度信息,使重建的光场图像保持更好的细节信息及几何一致性。

## 2 原理

### 2.1 光场成像及光场的 3D 对极平面图像

微透镜阵列光场相机的成像原理如图 1 所示,将微透镜阵列置于主透镜的焦平面处,微透镜可以将同一场景点不同方向的光线信息解耦,并记录在对应位置的传感器上,从而得到光线的空间信息和角度信息<sup>[21]</sup>。每个微透镜捕获一个场景点不同方向的光线信息,一个方向代表一个视角,视角的总数即为光场图像的角度分辨率。光场图像的空间分辨率是指捕获的场景点的数量,即微透镜的数量。在所有微透镜中固定角度坐标,将对应的像素点全部提取,可得到光场图像的一个视角图像,即子孔径图像。

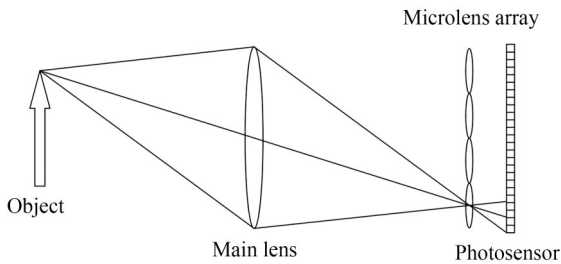


图 1 微透镜阵列光场相机成像原理示意图<sup>[21]</sup>

Fig. 1 Schematic of imaging principle for microlens array light field camera

如图 2 所示,在光场图像  $L(x, y, s, t)$  中,  $x$  和  $y$  表示二维空间坐标,  $s$  和  $t$  表示二维角度坐标。在水平方向,通过固定  $y$  轴坐标和  $t$  轴坐标,将  $x$  轴和  $s$  轴的像素点都提取出来,即可得到水平 EPI  $E_{y,t}(x, s)$ 。只固定角度坐标  $t$ ,将  $s$  轴的子孔径

图像进行堆叠,即可得到水平 3D EPI  $I_x(x, y, s)$ 。在垂直方向执行类似的操作即可得到垂直 EPI  $E_{x,s}(y, t)$  和垂直 3D EPI  $I_y(x, y, t)$ <sup>[20]</sup>。EPI 清晰地反映了水平方向或垂直方向上每个场景点在各个视角图像中的移动轨迹,能够体现光场图像的几何一致性。3D EPI 包含光场图像的二维空间信息和一维角度信息, EPI 包含光场图像的一维空间信息和一维角度信息。

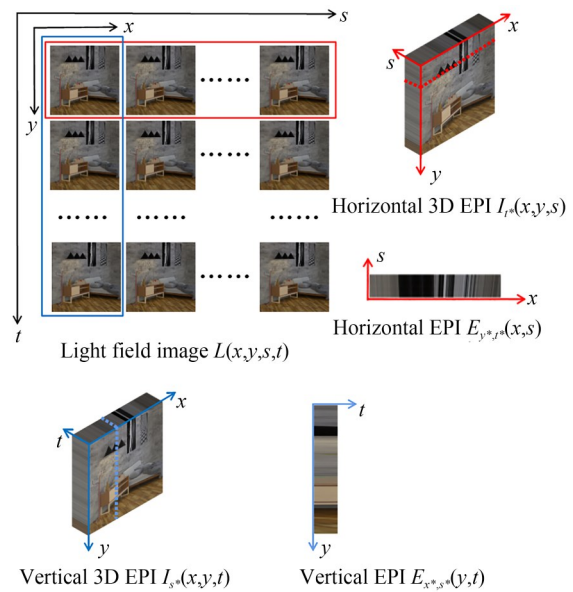


图 2 光场图像及其 3D EPI 与 EPI

Fig. 2 Light field image and its 3D EPI and EPI

### 2.2 方法框架

对于 4D 光场图像  $L(x, y, s, t)$ <sup>[11]</sup>,为了简化数学表达式,二维空间坐标  $(x, y)$  用  $x$  来简化表示,二维角度坐标  $(s, t)$  将用  $s$  来简化表示,即将  $L(x, y, s, t)$  简化表示为  $L(x, s)$ 。光场图像角度超分辨重建利用稀疏光场图像  $L'(x, s')$  重建密集光场图像  $\hat{L}(x, s)$ ,该问题可表示为:

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

其中  $f$  为光场图像角度超分辨重建对应的映射函数。本文提取密集光场图像 4 个角的子孔径图像作为稀疏光场图像,即输入图像。结合光场图像的 3D EPI 包含空间信息和角度信息的特点,提出一种融合 3D EPI 的光场角度超分辨重建方法。该方法遵循基于深度信息的重建方法的一

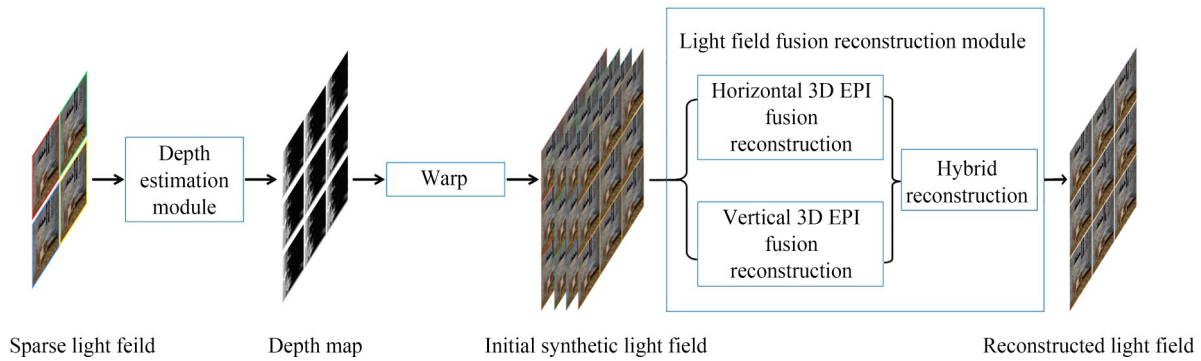


图3 融合3D EPI的光场图像角度超分辨重建网络框架

Fig. 3 Architecture of light-field angular super-resolution reconstruction via fusing 3D EPIs

般步骤,即深度估计、图像映射和光场融合重建,深度学习网络框架如图3所示。深度估计模块利用输入图像估计出密集光场所有视角的深度图。这些深度图可以把输入图像映射到每一个新视角位置,以合成初始光场。由于遮挡和深度图的不精确性,映射后的图像会产生几何扭曲。因此在最后的光场融合部分,与一般的直接融合映射图像的方法不同,本文的光场融合重建从水平和垂直两个方向分别融合由不同的输入图像映射产生的初始光场的3D EPI,再将水平重建结果与垂直重建结果混合,获得最终的高角度分辨率光场图像。

### 2.3 深度估计模块

光场图像具有规则的几何一致性,视差与深度成反比,这为深度估计模块无监督地估计所有视角的深度图提供了基础。给定位置为 $s_0$ 的已知视角图像 $L(x, s_0)$ , $d_{s_1}(x)$ 是位置为 $s_1$ 的待重建新视角图像 $L(x, s_1)$ 的深度图,则<sup>[12]</sup>:

$$L(x, s_1) = L(x + (s_1 - s_0)d_{s_1}(x), s_0). \quad (2)$$

深度估计模块利用输入图像估计出密集光场所有视角的深度图 $D(x, s)$ ,即:

$$D(x, s) = f_d(L'(x, s')), \quad (3)$$

其中 $f_d$ 为输入图像与深度图之间的映射关系。

深度估计模块在每个新视角位置生成对应景物的深度图,利用深度图将输入图像映射到新视角位置可以引入更多的场景信息,有利于减少大基线光场重建的伪影,提高光场重建的质量。光场图像的多视角结构决定了它具有多个方向的视差。为了更好地估计输入图像与深度图之

间的映射关系,本文将输入图像分为水平、垂直和对角3类图像对(如图4所示),分别送入结构相同的网络分支进行特征提取,学习不同方向的视差信息。稀疏光场本身就是密集光场的欠采样,光场图像4个角的子孔径图像作为整个场景的边缘视角,每个视角都包含重要的场景信息,特别是各个物体间的遮挡关系。所以,与Gul等<sup>[18]</sup>根据合成图像的位置选取距离最近的3幅输入图像组成3个图像对不同,本文利用4幅输入图像组成6个图像对,充分利用所有输入图像的视差信息。

深度估计模块首先通过特征提取网络对不同方向的图像对分别提取像素匹配特征。视差具有方向性,这样按视差方向进行特征提取的方法可以提取到更加精准的匹配特征,这些特征输入到后续的深度估计网络来生成深度图。如图4所示,特征提取网络分支类似于Epinet<sup>[22]</sup>,图像对首先通过1个包含32个 $3 \times 3$ 卷积核的卷积层(Conv),然后连续通过3个结构为“Conv-ReLU-Conv”的残差块(ResBlock),其中残差块的输入作为残差部分与输出相加,且残差块中的卷积层同样由32个 $3 \times 3$ 卷积核组成。特征提取完毕后,将6个分支提取的特征按顺序串联到一起作为深度估计网络的输入。深度估计网络由7个卷积层组成,其中前2个卷积层使用扩张率为2的 $5 \times 5$ 卷积,其余为 $3 \times 3$ 卷积层。深度估计网络除了最后一个卷积层外,其余每个卷积层后都跟一个ReLU激活层。

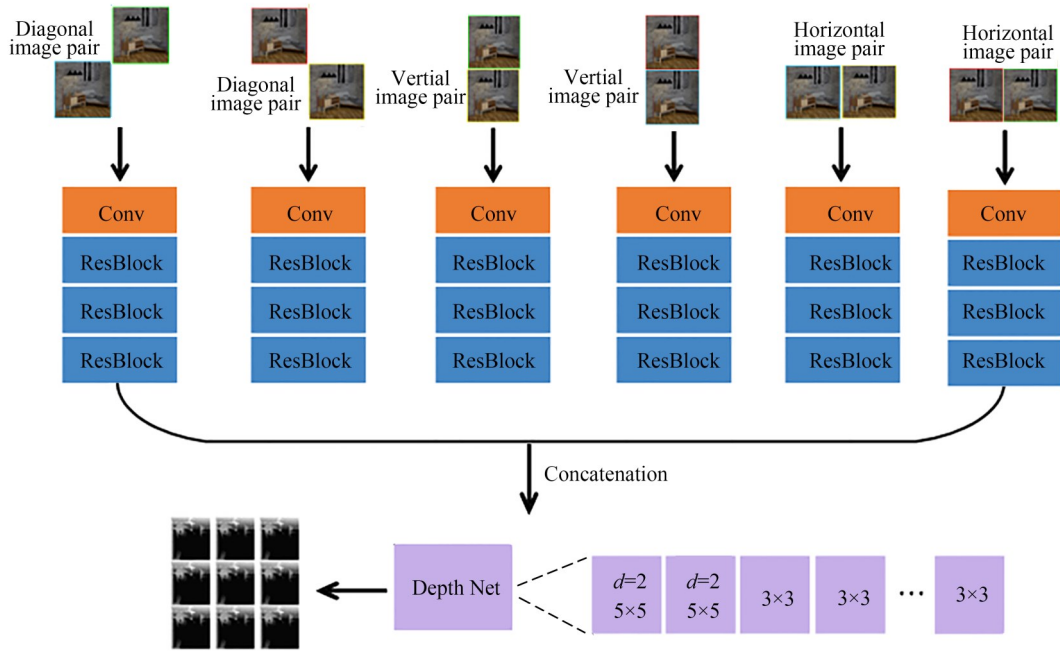


图 4 深度估计模块

Fig. 4 Depth estimation module

2.4 图像映射

图像映射利用深度估计模块得到的深度图将输入图像映射到每个新视角位置。将位置为  $s'$  的输入图像  $L'(x, s')$  映射到位置为  $s$  的新视角, 得到映射图像  $W_s(x, s)$ , 这一过程可以表示为:

$$W_s(x, s) = f_w(L'(x, s'), D(x, s)) = L'(x + (s - s')D(x, s), s'), \quad (4)$$

其中  $f_w$  表示图像映射过程。

由于有 4 幅输入图像, 每幅输入图像被分别映射到所有新视角, 所以每个新视角会有 4 幅映射图像, 即形成 4 个初始合成光场, 如图 5 所示。

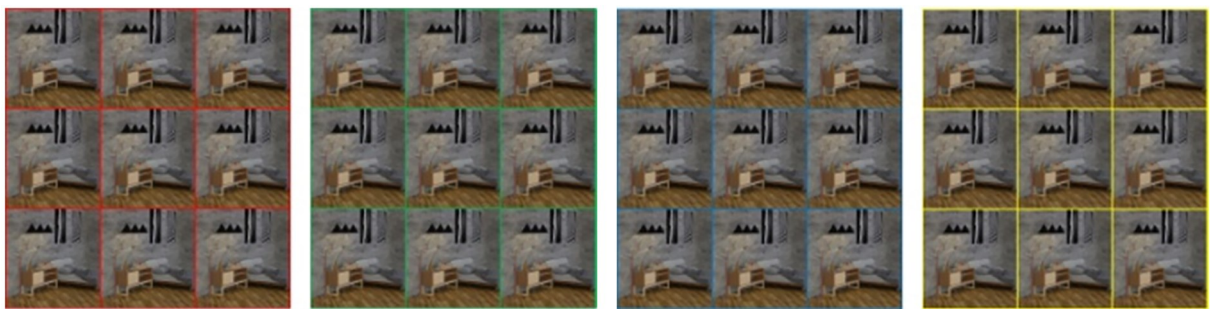


图 5 四个初始合成光场

Fig. 5 Four initial synthetic light fields

2.5 光场融合重建模块

光场融合重建模块利用 4 个初始合成光场生成最终的高角度分辨率光场图像。与通常的直接融合所有映射图像或构建修复网络不同, 本文通过评估 4 个初始合成光场的 3D EPI 来引导光场融合。光场融合重建步骤旨在从映射图像中

获取正确的光场信息, 这不仅要求每个视角图像保持正确的空间信息, 还要保证各视角间的几何关系。Navarro 等<sup>[17]</sup>在每个新视角位置对来自不同输入图像的映射图像的空间域进行卷积, 确定不同映射图像的融合权重, 以保留正确的空间信息。Jin 等<sup>[16]</sup>考虑到不同视角之间的角度相关性,

在对初始光场图像的空间域卷积之后又增加了对角度域的卷积。但是光场图像的空间信息和角度信息不是独立的两部分,它们之间存在密切的关联,这个关联体现在EPI中反映物体移动轨迹的斜线上。考虑空间信息和角度信息可以使重建光场保持更好的几何一致性,所以可以通过计算每个初始光场的EPI对最终光场的EPI的贡献来融合初始光场,而3D EPI比EPI多了一维空间信息,有助于重建光场保持更好的细节信息,所以最终通过融合4个初始合成光场的3D EPI来重建光场图像。

光场融合重建模块首先对4个初始合成光场进行水平3D EPI融合重建和垂直3D EPI融合重建,再将水平重建结果和垂直重建结果进行混合重建,得到最终的光场图像。这一过程可以表示为:

$$\hat{L}(x, s) = f_b(f_h(W_s(x, s)), f_v(W_s(x, s))), \quad (5)$$

其中:  $W_s(x, s)$  表示映射图像,即初始合成光场;  $f_h$  表示水平3D EPI融合重建,  $f_v$  表示垂直3D EPI融合重建,  $f_b$  表示水平垂直混合重建。对4个初始合成光场的角度坐标为  $t=1$  的子孔径图像分别堆叠,得到如图6所示的4个水平3D EPI。将



图6 4个初始合成光场的角度坐标  $t=1$  时的水平3D EPI  
Fig. 6 Horizontal 3D EPI at angular coordinate  $t=1$  of four initial synthetic light fields

这4个水平3D EPI分别输入3D EPI评估网络,计算出各自对重建光场图像的角度坐标为  $t=1$  的水平3D EPI的贡献,得到它们的融合权重,并进行加权融合,完成角度坐标为  $t=1$  的水平重建。同理,对其余的水平3D EPI进行融合,得到水平3D EPI融合重建的光场图像  $\hat{L}_h(x, s)$ ;对垂直3D EPI执行上述操作,得到垂直3D EPI融合重建的光场图像  $\hat{L}_v(x, s)$ 。最后,将  $\hat{L}_h(x, s)$  和  $\hat{L}_v(x, s)$  输入2个3D卷积层进行混合重建,得到最终的重建光场图像。其中,2个3D卷积层的卷积核尺寸分别为  $(5, 3, 3)$  和  $(3, 3, 3)$ ,步长分别为  $(4, 1, 1)$  和  $(1, 1, 1)$ 。

在每个新视角位置,若直接对4幅映射图像进行融合,那么每个像素点只有一个评价价值,且只在空间域进行融合权重的计算。本文提出的3D EPI融合方法通过水平重建和垂直重建,使每个像素点有两个评价价值,可以更好地保持细节信息,并且结合空间域和角度域来计算融合权重,有利于维持光场图像的几何一致性。

受 Wu 等<sup>[23]</sup>的启发,3D EPI评估网络使用带跳跃连接的编码器-解码器结构,如图7所示。编码器由3个卷积-池化层组成,每一个池化层都将图像的分辨率降为原来的二分之一。解码器由3个反卷积层组成,每个反卷积层都与对应的卷积层的高分辨率特征相连接,使解码器能够更好地捕捉细节信息。3D EPI评估网络的输入是角度维度为7的3D EPI,输出是对应的7个评分图。对于4个初始合成光场的角度坐标为  $t=1$  的水平3D EPI,分别进行计算评分后,每个视角的4

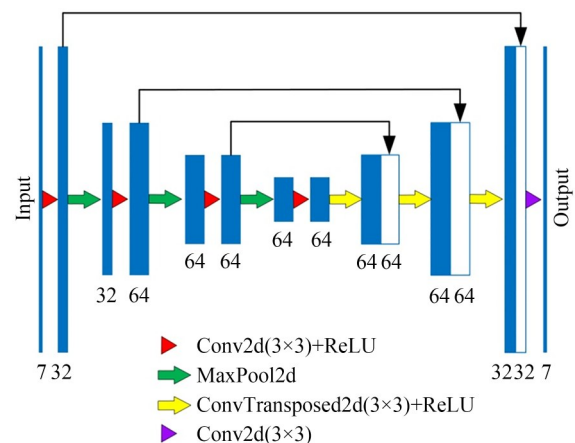


图7 三维EPI评估网络

Fig. 7 Three-dimensional EPI evaluation network

幅映射图像都有对应的评分图。接着通过 Softmax 函数将每个视角位置的 4 个评分图映射到  $[0,1]$  之间,得到 4 幅映射图像的融合权重,并进行加权融合。对其余的每组 3D EPI 都执行相同的操作。

## 2.6 损失函数

与 Jin 等<sup>[16]</sup>类似,本文提出的光场角度超分辨重建方法的损失函数  $l$  定义为:

$$l = l_d + l_b + \lambda l_e, \quad (6)$$

其中  $l_d$  为深度估计损失函数。因为真实的光场深度图很难获取,所以通过度量初始合成光场与真实光场之间的差异来优化深度图,并且通过惩罚空间梯度对深度图进行平滑约束。 $l_d$  定义为:

$$l_e = \sum_{y,t} \left( \left| \nabla_x E_{y,t}(x,s) - \nabla_x \hat{E}_{y,t}(x,s) \right| + \left| \nabla_s E_{y,t}(x,s) - \nabla_s \hat{E}_{y,t}(x,s) \right| \right) + \sum_{x,s} \left( \left| \nabla_y E_{x,s}(y,t) - \nabla_y \hat{E}_{x,s}(y,t) \right| + \left| \nabla_t E_{x,s}(y,t) - \nabla_t \hat{E}_{x,s}(y,t) \right| \right), \quad (9)$$

其中:  $E_{y,t}(x,s)$  和  $E_{x,s}(y,t)$  分别表示真实光场的水平 EPI 和垂直 EPI,  $\hat{E}_{y,t}(x,s)$  和  $\hat{E}_{x,s}(y,t)$  分别表示重建光场的水平 EPI 和垂直 EPI,  $\lambda$  为  $l_e$  的权重,本文设置为 1。

## 3 实验结果及分析

实验使用的训练集为 HCI new<sup>[24]</sup> 的 20 个光场图像,每个光场图像的角度分辨率为  $9 \times 9$ ,空间分辨率为  $512 \times 512$ 。针对角度分辨率为  $2 \times 2$  的稀疏光场图像进行角度超分辨重建,生成角度分辨率为  $7 \times 7$  的密集光场图像。在制作训练集时,将  $9 \times 9$  光场周围一圈的视角剪裁掉,仅使用中间  $7 \times 7$  的子孔径图像,并且每个子孔径图像的空间分辨率被随机剪裁为  $96 \times 96$ ,即使用  $7 \times 7 \times 96 \times 96$  的光场图像作为密集光场图像。实验使用密集光场图像的 4 个角度坐标为  $(0,0)$ ,  $(0,6)$ ,  $(6,0)$ ,  $(6,6)$  的子孔径图像作为输入图像。

本文使用 PyTorch 框架,在 NVIDIA GeForce RTX 3050 上对提出的模型进行训练。实验使用 Adam 优化器 ( $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ) 来训练网络,初始学习率设置为 0.0001,且每 500 个周期衰减 0.5 倍。

### 3.1 对比实验

为了初步验证本文方法是否能够提高光场

$$l_d = \sum_{x,s} \left( \sum_{s'} |L(x,s) - W_{s'}(x,s)| + \nabla_x D(x,s) \right), \quad (7)$$

其中:  $L(x,s)$  代表真实的高角度分辨率光场图像,  $\nabla_x D(x,s)$  是深度图  $D(x,s)$  的空间梯度。 $l_b$  为融合重建损失函数,用于度量重建光场图像与真实光场图像的相似度,定义为:

$$l_b = \sum_{x,s} |L(x,s) - \hat{L}(x,s)|, \quad (8)$$

其中  $\hat{L}(x,s)$  为重建的高角度分辨率光场图像。 $l_e$  为 Jin 等<sup>[16]</sup>提出的 EPI 梯度损失函数 (EPI gradient loss),通过计算重建光场图像的 EPI 梯度和真实光场图像的 EPI 梯度之间的 L1 距离来保持光场图像的几何一致性,梯度在 EPI 的空间维度和角度维度上计算。 $l_e$  定义为:

角度超分辨重建的质量,在 HCI new<sup>[24]</sup> 的测试集上对本文的重建结果与 Yun 等<sup>[19]</sup> 的重建结果进行比较,通过峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 与结构相似性 (Structural Similarity, SSIM) 来定量评估光场的重建质量。实验结果如表 1 所示,可以看出本文方法的重建结果的 PSNR 和 SSIM 都较高,说明该方法能够提高光场角度超分辨重建的质量。

表 1 光场角度超分辨重建结果的比较

Tab.1 Comparison of angular super-resolution reconstruction results on light field images

HCI new	PSNR/dB	SSIM
Yun	33.80	0.934
Ours	34.44	0.942

为了进一步分析本文方法的有效性及其适用场景,选取 lf-syn<sup>[14]</sup>, RCA-LF<sup>[11]</sup>, LFASR<sup>[16]</sup> 3 种方法进行对比实验。在对比实验中,除了 RCA-LF<sup>[11]</sup>,其他方法都是基于深度信息的重建方法。公平起见,各个方法均使用相同的数据集重新训练,且这些方法均使用各自论文提供的代码和训练模型。对比实验分别在合成光场图像和真实光场图像上进行。

### 3.1.1 合成光场图像

对于合成光场图像,实验使用 HCI new<sup>[24]</sup>中的 bedroom, bicycle, herbs, dishes 和 HCI old<sup>[25]</sup>中的 Buddha, Mona 来评估各个方法的重建结果。所有测试子孔径图像在评估时都被剪裁掉四周的 22 pixels。图 8 展示了各方法在场景 bedroom, bicycle 上的角度超分辨重建结果的中心视角图像及其局部放大,也展示了局部放大部分的中心水平 EPI。从图中可以看出,lf-syn<sup>[14]</sup>的重建结果整体较为模糊,主要是因为该方法仅用相邻的子孔径图像来生成中间图像,利用的光场信息较

少。该方法还十分耗时,重建一个场景大概需要 3 h。RCA-LF<sup>[11]</sup>没有利用深度信息,因此重建结果的整体性较差,如图 8(b)中 EPI 的斜线有较多的错误。LFASR<sup>[16]</sup>通过 EPI 损失函数约束重建结果的几何一致性,重建结果的整体性更好,但是一些复杂的纹理部分会出现模糊。本文提出的角度超分辨方法在水平方向和垂直方向分别对 3D EPI 进行整体融合,能够更好地保持纹理细节信息,且重建光场的几何一致性较好。表 2 从定量角度展示了各方法的重建结果,可以看出,本文方法在大部分场景下表现出更好的性能。

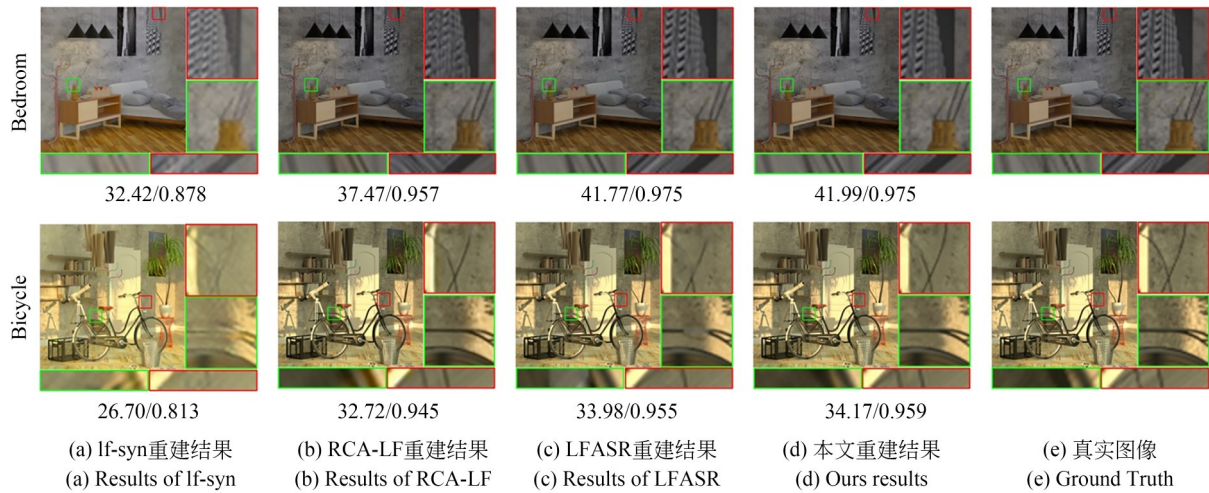


图 8 合成光场图像角度超分辨重建结果

Fig. 8 Angular super-resolution reconstruction results of synthetic light field images

表 2 合成光场图像上角度超分辨重建结果的比较

Tab. 2 Comparison of angular super-resolution reconstruction results on synthetic light field images

Method	HCI new	Buddha	Mona
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
lf-syn	27.56/0.800	35.09/0.935	32.87/0.892
RCA-LF	32.06/0.899	44.48/0.989	45.19/0.989
LFASR	34.39/0.937	45.78/0.991	45.88/0.990
Ours	34.44/0.942	46.13/0.992	45.58/0.990

### 3.1.2 真实光场图像

对于真实光场图像,实验选取 Stanford Lytro Light Field<sup>[26]</sup>的 Bikes 和 Kalantari 等<sup>[14]</sup>的 Rock 和 IMG\_1528\_eslf。在制作测试集时,先将这些光场的子孔径图像尺寸统一为  $540 \times 360$ ,为了

更好地分析结果,所有测试光场的子孔径图像在评估时都被剪裁掉四周的 22 pixel。图 9 展示了各方法在场景 bikes\_20\_eslf, IMG\_1528\_eslf 上的角度超分辨重建结果的中心视角图像及其局部放大,也展示了局部放大部分的中心水平 EPI。真实光场图像往往有着更复杂的遮挡关系,这更加考验重建方法对深度关系的计算。lf-syn<sup>[14]</sup>在深度估计阶段是手动提取特征,生成的深度图不精确,重建场景有很多重影。RCA-LF<sup>[11]</sup>忽略几何信息,直接对角度维度进行上采样,导致遮挡边缘出现模糊。相比于 LFASR<sup>[16]</sup>,本文方法从水平、垂直、对角方向分别提取视差特征,一定程度上提高了深度图的精度,在主观视觉上表现出更好的遮挡关系。表 3 展示了各方法重建结果的 PSNR 和 SSIM,可以看出,本文方法在真实光场图像上的重建效果优于其他方法。

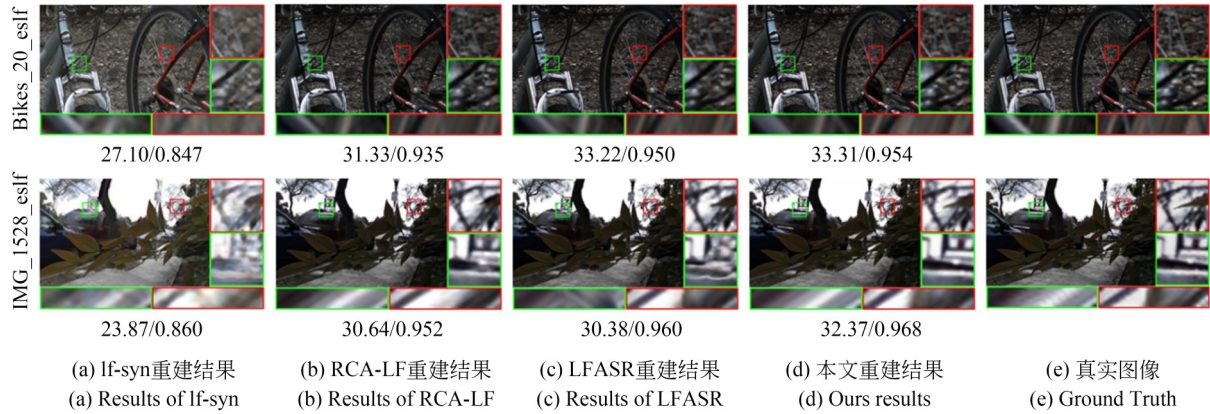


图 9 真实光场图像的角度超分辨重建结果

Fig. 9 Angular super-resolution reconstruction results of real-world light field images

表 3 真实光场图像上角度超分辨重建结果的比较

Tab. 3 Comparison of angular super-resolution reconstruction results on real-world light field images

Method	Bikes	IMG_1528_eslf	Rock
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
If-syn	29.31/0.904	23.87/0.860	27.35/0.892
RCA-LF	34.02/0.965	30.64/0.952	34.97/0.972
LFASR	36.35/0.974	30.38/0.960	38.53/0.981
Ours	36.73/0.975	32.37/0.968	38.72/0.982

### 3.2 消融实验

为了验证水平 3D EPI 融合重建分支和垂直 3D EPI 融合重建分支的组合能否充分利用光场的高维信息,进而提高重建光场图像的质量,对这两个分支进行了消融实验,其他部分保持不变。消融实验(1):只保留水平 3D EPI 融合重建

分支;消融实验(2):只保留垂直 3D EPI 融合重建分支。实验结果如表 4 所示,只进行水平重建或垂直重建时的效果比较接近,这表明水平方向上的三维信息和垂直方向上的三维信息同等重要。本文方法将两者组合起来,充分利用两个方向上的高维信息,重建效果比单独使用水平重建分支或单独使用垂直重建分支的效果好。

为了验证深度估计模块按不同视差方向分别提取特征能够有效地提高深度图的精度,还设计了消融实验(3),即去掉特征提取部分,直接将输入图像输入深度估计网络,深度估计结果如表 4 所示。从表 4 可以看出,去掉特征提取部分后的重建效果比单分支融合重建还差,证明了按视差方向进行特征提取有利于提高深度估计的准确性,进而提高光场图像的重建质量。

表 4 消融实验的角度超分辨重建结果

Tab. 4 Angular super-resolution reconstruction results produced in ablation experiments

LF images	Buddha	Mona	IMG_1528_eslf	Rock
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Ablation (1)	45.97/0.992	45.30/0.990	32.06/0.967	38.45/0.981
Ablation (2)	45.97/0.992	45.19/0.990	31.92/0.967	38.48/0.981
Ablation (3)	45.36/0.990	45.27/0.990	30.11/0.957	37.97/0.980
Ours	46.13/0.992	45.58/0.990	32.37/0.968	38.72/0.982

### 3.3 光场应用

重建光场具有高角度分辨率,包含更多的场景信息,可以进行深度估计、重聚焦等光场应用。这里采用 Epinet<sup>[22]</sup>的方法,利用本文方法的重建

结果对光场图像进行深度估计,结果如图 10 所示。从图中可以看出,利用本文的重建结果能够估计出场景的深度图,且边缘部分能够保持清晰的结构。

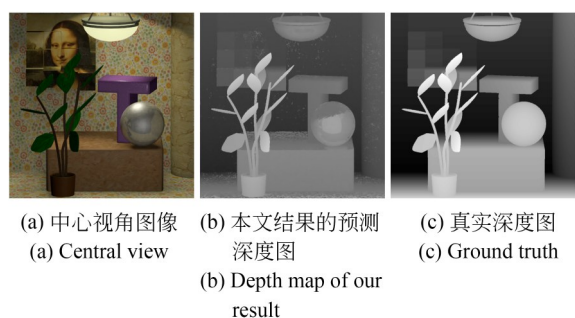


图 10 利用本文重建结果预测的深度图

Fig. 10 Depth map predicted by reconstruction results in this article

## 4 结 论

本文针对利用稀疏光场图像重建密集光场图像的问题,提出了融合 3D EPI 的光场角度超分辨率重建方法。该方法通过对输入图像提取不同方向的视差特征,有效地提高了深度估计的准

确性,用生成的深度图将输入图像映射到新视角位置,形成初始合成光场,最后从水平方向和垂直方向分别对初始合成光场的 3D EPI 进行加权融合,再将两个方向的融合结果进行混合重建,得到最终的高质量重建光场图像。本文方法适用于用微透镜阵列光场相机采样的大基线稀疏光场图像,可以提高光场角度超分辨率重建的质量,在合成光场数据集和真实光场数据集上的重建效果均得到了提高,PSNR 值的提升幅度最高达 1.99%,更好地保持了重建光场图像的细节信息及几何一致性,在遮挡区域展现出更好的重建效果。但是,光场图像的采样方式不止一种,对于用相机阵列采样的光场图像,其各视角之间的基线会更宽,本文方法的重建质量就较低。未来会继续深入探索光场图像包含的丰富信息,优化本文方法,使该方法可以适用于相机阵列采样的光场图像。

## 参考文献:

- [1] LEVOY M, HANRAHAN P. Light field rendering [C]. *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. New York: ACM, 1996: 31-42.
- [2] 刘雪岩,许聿达,雷建昕,等. 基于视差放大与超分辨率的三维光场腹腔镜标定[J]. *光学精密工程*, 2022, 30(5):510-517.  
LIU X Y, XU Y D, LEI J X, *et al.* Three-dimensional light field endoscope calibration based on light field disparity amplifier and super-resolution network [J]. *Optics and Precision Engineering*, 2022, 30(5):510-517. (in Chinese)
- [3] 池汉彬,段辉高,胡跃强. 超构表面在三维成像与显示技术中的应用[J]. *光学精密工程*, 2022, 30(15):1775-1801.  
CHI H B, DUAN H G, HU Y Q. Application of metasurfaces in three-dimensional imaging and display [J]. *Optics and Precision Engineering*, 2022, 30(15):1775-1801. (in Chinese)
- [4] 吕晓波,刘宇丰,李毅威,等. 快照式光谱光场成像技术[J]. *光学精密工程*, 2021, 29(2):220-230.  
LÜ X B, LIU Y F, LI Y W, *et al.* Snapshot spectral light-field imaging technology [J]. *Optics and Precision Engineering*, 2021, 29(2): 220-230. (in Chinese)
- [5] 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): 1-13.
- [6] 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.
- [7] 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)*. December 7-13, 2015, Santiago, Chile. IEEE, 2016: 57-65.
- [8] 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, USA. IEEE, 2017: 1638-1646.
- [9] 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.
- [10] WANG Y L, LIU F, ZHANG K B, *et al.* High-fidelity view synthesis for light field imaging with extended pseudo 4DCNN [J]. *IEEE Transactions on Computational Imaging*, 2020, 6: 830-842.

- [11] SALEM A, IBRAHEM H, KANG H S. RCA-LF: dense light field reconstruction using residual channel attention networks[J]. *Sensors*, 2022, 22(14): 5254.
- [12] 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.
- [13] PENNER E, ZHANG L. Soft 3D reconstruction for view synthesis [J]. *ACM Transactions on Graphics*, 2017, 36(6): 1-11.
- [14] KALANTARI N K, WANG T C, RAMAMOORTHY R. Learning-based view synthesis for light field cameras [J]. *ACM Transactions on Graphics*, 2016, 35(6): 1-10.
- [15] VADATHYA A K, GIRISH S, MITRA K. A deep learning framework for light field reconstruction from focus-defocus pair: a minimal hardware approach [C]. *Computational Cameras and Displays Workshop*. 22, 2018, Salt Lake City, UT, USA. IEEE, 2018.
- [16] 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.
- [17] NAVARRO J, SABATER N. Learning occlusion-aware view synthesis for light fields[J]. *Pattern Analysis and Applications*, 2021, 24(3): 1319-1334.
- [18] GUL M S K, MUKATI M U, BÄTZ M, *et al.* Light-field view synthesis using A convolutional block attention module[C]. *2021 IEEE International Conference on Image Processing (ICIP)*. September 19-22, 2021, Anchorage, AK, USA. IEEE, 2021: 3398-3402.
- [19] YUN S, JANG J, PAIK J. Geometry-aware light field angular super resolution using multiple receptive field network[C]. *2022 International Conference on Electronics, Information, and Communication (ICEIC)*. February 6-9, 2022, Jeju, Korea, Republic of. IEEE, 2022: 1-3.
- [20] WANNER S, GOLDLUECKE B. Globally consistent depth labeling of 4D light fields [C]. *2012 IEEE Conference on Computer Vision and Pattern Recognition*. June 16-21, 2012, Providence, RI, USA. IEEE, 2012: 41-48.
- [21] NG R, LEVOY M, BRÉDIF M, *et al.* *Light Field Photography with a Hand-held Plenoptic Camera* [D]. California: Stanford University, 2005.
- [22] 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*. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 4748-4757.
- [23] 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.
- [24] HONAUER K, JOHANNSEN O, KONDERMANN D, *et al.* *A Dataset and Evaluation Methodology for Depth Estimation on 4D Light Fields* [M]. Computer Vision-ACCV 2016. Cham: Springer International Publishing, 2017: 19-34.
- [25] WANNER S, MEISTER S, GOLDLUECKE B. Datasets and benchmarks for densely sampled 4D light fields [C]. *18th International Workshop on Vision, Modeling, and Visualization*. 11-13, 2013, Lugano, Switzerland. The Eurographics Association, 2013, 13: 225-226.
- [26] RAJ A S, LOWNEY M, SHAH R, *et al.* *Stanford Lytro Light Field Archive*[EB/OL]. (2016-10) [2022-12-29]. <http://lightfields.stanford.edu/LF2016.html>.

#### 作者简介:



陈纯毅(1981—),男,重庆人,博士,教授,博士生导师,2003年、2006年、2009年于长春理工大学分别获得学士、硕士和博士学位,主要从事三维显示与光场图像处理、虚拟现实、光电系统仿真等方面的研究。E-mail: chenchunyi@hotmail.com



范晓辉(1998—),女,河南南阳人,硕士研究生,2020年于河南科技学院获得学士学位,主要研究方向为计算机图形学。E-mail: 963014193@qq.com