

# 联合遮挡约束与残差补偿的特征金字塔光流计算方法

范兵兵<sup>1</sup>, 何庭建<sup>1</sup>, 张聪炫<sup>1,2</sup>, 陈震<sup>1</sup>, 黎明<sup>1</sup>

(1. 南昌航空大学无损检测技术教育部重点实验室, 江西南昌 330063; 2. 中国科学院自动化研究所, 北京 100190)

**摘要:** 针对现有深度学习光流计算模型在运动遮挡和大位移等场景下光流计算的准确性与鲁棒性问题, 本文提出一种联合遮挡约束与残差补偿的特征金字塔光流计算方法. 首先, 构造基于遮挡掩模的光流约束模块, 通过预测遮挡掩模特征图抑制变形特征的边缘伪影, 克服运动遮挡区域的图像边缘模糊问题. 然后, 采用特征图变形策略构建基于特征变形的光流残差补偿模块, 利用该模块学习到的残差光流细化原始光流场, 改善大位移运动区域的光流计算效果. 最后, 采用特征金字塔框架构建联合遮挡约束与残差补偿的光流计算网络模型, 提升大位移和运动遮挡场景下的光流计算精度. 分别采用 MPI-Sintel (Max-Planck Institute and Sintel) 和 KITTI (Karlsruhe Institute of Technology and Toyota Technological Institute) 数据集对本文方法和代表性传统光流计算方法、深度学习光流计算方法进行综合对比分析, 实验结果表明本文方法相对于其他方法能够有效提升大位移和运动遮挡场景下的光流计算精度与鲁棒性.

**关键词:** 光流; 遮挡约束; 残差补偿; 特征金字塔网络; 深度学习; 边缘保护

**基金项目:** 国家自然科学基金(No.61866026, No.61866025); 江西省优势科技创新团队计划(No.20165BCB19007); 江西省技术创新引导类计划项目(No.20212AEI91005); 江西省自然科学基金重点项目(No.20202ACB214007); 航空科学基金(No.2018ZC56008); 中国博士后科学基金(No.2019M650894); 江西省教育厅科学技术研究项目(No.GJJ210910); 江西省图像处理与模式识别重点实验室开放基金资助(No.ET202104413)

中图分类号: TP391

文献标识码: A

文章编号: 0372-2112(2023)03-0648-10

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20211052

## Feature Pyramid Optical Flow Estimation Method Jointing Occlusion Constraint and Residual Compensation

FAN Bing-bing<sup>1</sup>, HE Ting-jian<sup>1</sup>, ZHANG Cong-xuan<sup>1,2</sup>, CHEN Zhen<sup>1</sup>, LI Ming<sup>1</sup>

(1. Key Laboratory of Nondestructive Testing, Ministry of Education, Nanchang Hangkong University, Nanchang, Jiangxi 330063, China;

2. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)

**Abstract:** To improve the accuracy and robustness of the deep-learning based optical flow estimation under motion occlusions and large displacements, we propose a feature pyramid optical flow computation method by jointing the occlusion constraint with residual compensation. First, an optical flow constraint module is designed based on the learning occlusion mask. The proposed constraint module predicts the occlusion feature map to restrain the edge artifacts of the warping features, which is able to overcome the issue of edge blurring in the motion occlusion areas. Second, a residual compensation module is constructed by using the feature map warping strategy, and the residual optical flows learned from the presented module are employed to refine the original flow fields. Third, the proposed occlusion constraint model and residual compensation module are incorporated into a feature pyramid framework to construct an optical flow estimation network. Finally, the MPI-Sintel (Max-Planck Institute and Sintel) and KITTI (Karlsruhe Institute of Technology and Toyota Technological Institute) datasets are employed to conduct a comprehensive comparison between the proposed method and the representative traditional optical flow methods, deep-learning optical flow methods. The experimental results demonstrate that the presented method significantly improves the accuracy and robustness of optical flow estimation under large displacements and motion occlusions.

**Key words:** optical flow; occlusion constraint; residual compensation; feature pyramid network; deep-learning; edge-

preserving

Foundation Item(s): National Natural Science Foundation of China (No.61866026, No.61866025); Advantage Subject Team of Jiangxi Province (No.20165BCB19007); Technological Innovation Guidance Program of Jiangxi Province (No.20212AEI91005); Natural Science Foundation of Jiangxi Province (No.20202ACB214007); Aeronautical Science Foundation of China (No.2018ZC56008); China Postdoctoral Science Foundation (No.2019M650894); Science and Technology Program of Education Department of Jiangxi Province (No.GJJ210910); Open Fund of Jiangxi Key Laboratory of Image Processing and Pattern Recognition (No.ET202104413)

## 1 引言

光流是指图像序列中运动目标或场景表面像素点的二维运动矢量,其不仅提供了图像中运动目标和场景的运动参数,还携带了丰富的三维结构信息.研究成果被广泛应用于表情识别<sup>[1]</sup>、姿态计算<sup>[2]</sup>、目标跟踪<sup>[3]</sup>、三维重建<sup>[4]</sup>等高级视觉任务.

根据研究路线与计算方法的不同,光流计算技术可大致分为传统光流计算方法和深度学习光流计算方法.传统光流计算方法主要以基于变分理论的光流计算方法为主,1981年Horn和Schunck<sup>[5]</sup>通过对图像亮度守恒假设添加正则化约束,提出了首个变分光流计算模型.随后,变分光流计算方法主要集中在能量泛函设计与优化策略等方面研究<sup>[6-11]</sup>.虽然传统变分光流计算方法的精度与可靠性已得到大幅提升,但由于该方法需要大量迭代计算最小化能量泛函,导致其时间消耗过大,难以应用于对实时性要求较高的高级视觉任务.

近年来,随着以卷积神经网络为代表的深度学习理论与方法的快速发展,基于深度学习的光流计算技术研究取得巨大进展<sup>[12]</sup>.Dosovitskiy等人<sup>[13]</sup>首次提出基于有监督学习的光流计算模型FlowNet,该模型利用U-Net网络直接从图像序列预测光流场.随后,Ilg等人<sup>[14]</sup>在FlowNet模型基础上通过堆叠多功能子网络构建了其升级版网络FlowNet2,显著提高了深度学习光流计算的精度.针对网络模型尺寸增大易于陷入过拟合问题,Ranjan等人<sup>[15]</sup>提出基于图像金字塔分层的光流计算网络模型,大幅减少了模型的参数量与模型复杂度,但由于网络模型的过于简化导致光流计算精度较低.为了在简化模型复杂度的同时保证光流计算精度,Sun等人<sup>[16]</sup>构造了基于特征金字塔的光流计算网络模型,通过捕捉图像序列上下文信息优化模块,既显著提高了大位移运动光流计算的精度又大幅降低模型参数量.鉴于卷积操作易导致光流边缘过度平滑的问题,Hui等人<sup>[17]</sup>提出特征匹配推理与局部正则化相结合的光流计算网络模型,有效保护了光流计算的运动边缘.针对运动遮挡区域光流计算的鲁棒性问题,Zhao等人<sup>[18]</sup>提出基于遮挡约束的非对称特征金字塔网络模型,通过遮挡信息去除无效遮挡区域,有效

抑制了遮挡边界的边缘模糊现象.此外,Lu等人<sup>[19]</sup>利用光流场变形特征匹配代价卷,进一步提高了运动边缘区域光流计算的精度.针对多维匹配代价卷计算量过大的问题,Yang等人<sup>[20]</sup>采用可分离四维卷积滤波处理匹配代价关系,在减少计算量的同时提高了光流计算的精度.

由于有监督光流计算方法依赖于大量标签数据参与模型训练,限制了该类方法在真实场景中的应用.借鉴变分光流能量泛函的设计思想,Jason等人<sup>[21]</sup>通过构造由数据项和平滑项组成的损失函数,首次实现了端到端的无监督学习光流计算.针对运动遮挡对光流预测的影响,Meister等人<sup>[22]</sup>首先利用前后向光流判别遮挡区域,然后提出无监督光流计算模型Unflow,有效提高了遮挡场景下无监督学习光流计算的鲁棒性.为了提高无监督学习光流计算的精度,Liu等人<sup>[23]</sup>构建了基于数据驱动的蒸馏学习无监督光流计算模型,利用Teacher网络的预测光流监督Student网络进行光流计算,有效提高了无监督学习光流预测的精度.为了充分利用有监督和无监督学习的优势,Lai等人<sup>[24]</sup>提出基于生成对抗网络的半监督光流计算方法,该方法无需对亮度守恒和运动平滑假设进行建模,而是利用生成对抗损失正则化光流计算,显著提高了小样本数据光流计算的准确性.随后,Song等人<sup>[25]</sup>提出基于残差补偿的半监督学习光流计算模型,通过残差网络由粗到细迭代细化光流,大幅提升了半监督学习光流模型的计算精度.虽然无监督与半监督学习方法无需大量标签数据训练网络模型,但由于网络模型学习的不充分,导致它们相对于有监督光流计算方法的计算精度仍存在较大差距.

虽然深度学习光流计算方法在计算精度和时效性方面已得到显著提高,但在运动遮挡与大位移场景下光流计算的准确性与鲁棒性仍然有待提高.针对该问题,本文提出一种联合遮挡约束与残差补偿的特征金字塔光流计算方法,通过遮挡掩模特征图抑制边缘伪影并利用残差光流细化原始光流场,有效提高了大位移和运动遮挡场景光流计算的精度与鲁棒性.

## 2 联合遮挡约束与残差补偿的光流计算模型

### 2.1 联合遮挡约束与残差补偿的特征金字塔光流计算网络

图1展示了本文提出的联合遮挡约束与残差补偿的特征金字塔光流计算网络模型,主要由特征金字塔编码网络和解码网络组成. 编码网络主要用于提取图像特征,解码网络用于光流计算. 运动遮挡约束模块和光流残差补偿模块均嵌入于特征金字塔解码网络,在光流计算过程中遮挡约束模块用于抑制光流的边缘模糊,残差补偿模块用于补偿大位移区域的光流信息.

如图1所示,将连续两帧图像输入特征金字塔编码

网络计算特征图,再将特征图输入解码网络计算光流. 在解码网络光流估计过程中,特征金字塔顶层计算出的光流经上采样后作为初始光流传递至下一层的遮挡掩模约束模块和残差补偿模块. 遮挡掩模约束模块首先根据当前层特征图和变形特征图计算遮挡掩模权重图,然后利用遮挡掩膜约束当前层变形特征,计算遮挡优化后的遮挡残差流,最后将遮挡残差流与初始光流叠加获得遮挡约束优化后的光流. 残差补偿模块首先利用遮挡优化后的光流再次变形当前层图像特征,然后计算大位移运动区域的二次残差光流,最后将遮挡优化后光流与二次残差光流叠加作为当前层解码网络计算的最终光流. 将当前层输出光流上采样后输入下一层解码器进行相同操作,直至金字塔解码网络最底层得到本文方法最终光流计算结果.

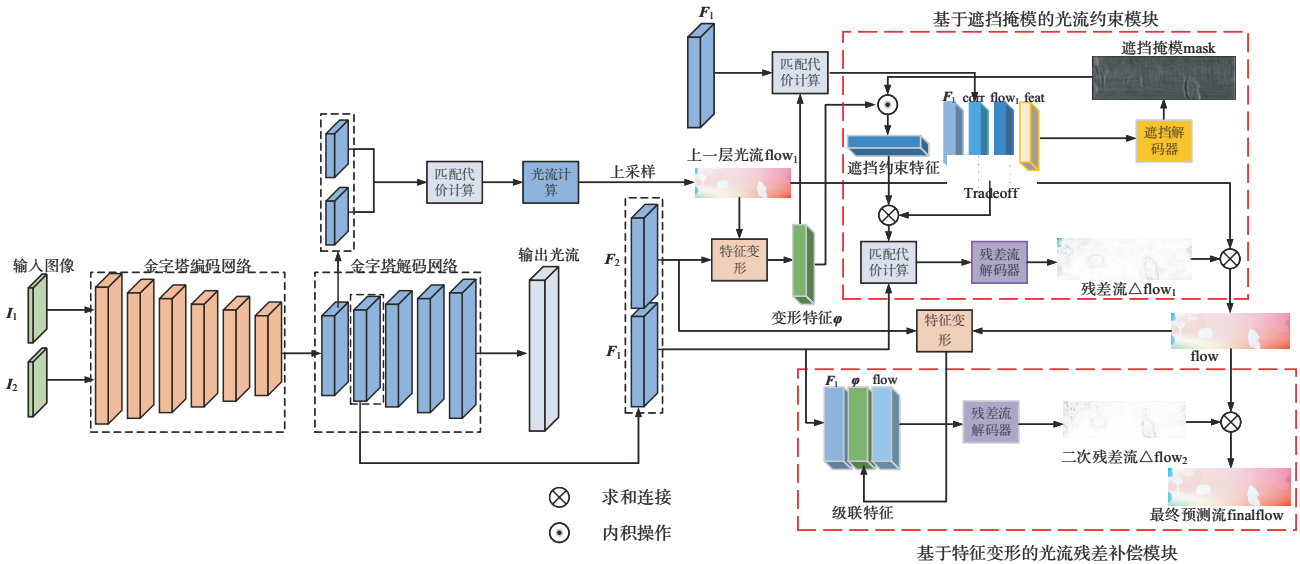


图1 联合遮挡约束和残差补偿的特征金字塔光流计算网络模型

### 2.2 基于遮挡掩模的光流约束模块

运动遮挡是指图像中运动物体或场景表面像素点在连续帧间的“时有时无”现象,通常发生在图像目标和运动平面的边缘区域. 由于光流的本质是描述图像像素点在连续帧图像间坐标变化的矢量,而遮挡区域的像素点在连续帧图像间并不对应,因此,光流计算在遮挡区域往往包含较大误差. 同时,为了实现特征之间的长程匹配,当前光流计算方法普遍采用变形策略解决该问题. 令  $I_1, I_2$  为连续两帧输入图像序列,  $w$  为计算出的位移矢量,表示  $I_1(X)$  与  $I_2(X+w)$  之间的对应关系,则变形策略可公式化如下:

$$(w \circ I_2)(X) \triangleq I_2(X + w(X)) \quad (1)$$

其中,  $X = (x, y)^T$  表示像素点坐标,  $\triangleq$  表示定义为:理想状态下,在所有非遮挡像素区域  $I_1(X) \approx (w \circ I_2)(X)$ . 然而,由于运动遮挡现象的存在,在变形过程中往往会引

入一些不可靠的信息,造成遮挡区域存在加倍或者伪影现象.

针对运动遮挡区域经特征变形后光流计算的边缘模糊问题,本文提出一种级联特征遮挡掩模预测模型,模型框架示意图如图2所示. 图中  $F_1$  和  $F_2$  分别表示特征金字塔编码网络提取到的参考帧与目标帧图像特征. 本文首先利用特征金字塔上一层光流与图像特征变形出目标帧图像特征,并计算参考帧特征与目标帧变形特征的匹配代价卷;然后将参考帧特征、匹配代价卷、上一层光流以及上一层反卷积后的级联特征在特征通道上进行特征级联;最后将级联特征输入遮挡解码器预测连续帧图像间的运动遮挡信息即遮挡掩模. 从图2可以看出,遮挡解码器包含四个卷积模块和一个卷积层,每个卷积模块都由一个卷积层(Convolution, Conv)、一个批量归一化处理层(Batch Normalization,

BN) 和一个激活函数层 (Rectified Linear Unit, ReLU) 组成. 其中, 卷积模块中的卷积层用于提取图像特征, 四个卷积模块输出特征图的通道依次递减, 分别为 128, 96, 64, 32; 批量归一化处理层用于对图像特征进行批量归一化处理, 促使图像特征分布相近以加快模型收敛; 激活函数层用于增加模型的非线性复杂度提高模型泛化性; 最后的卷积层用于预测遮挡掩模. 遮挡掩模是一种概率图, 其像素点灰度值的大小表示对应像素点被判定为遮挡像素点概率的大小, 像素点灰度值越大, 则该像素点被判定为是遮挡像素点的概率越大, 反之越小.

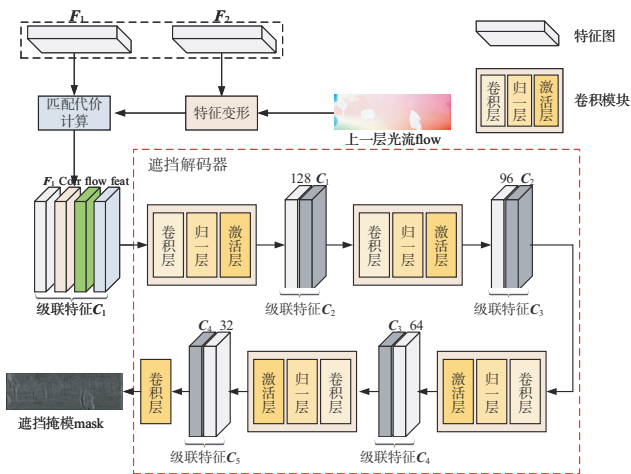


图2 基于级联特征的遮挡掩模预测模型

获取遮挡掩模后, 本文利用遮挡掩模特征图优化变形特征的边缘伪影, 以克服运动遮挡区域光流计算边缘模糊问题, 图3展示了基于遮挡掩模的光流约束优化过程. 首先利用上一级光流与目标特征进行特征变形; 然后将遮挡权重图与变形特征图执行内积操作, 以抑制变形特征图中的边缘伪影现象; 接着对特征信息权重项 (级联特征) 与经遮挡约束后的特征图执行加法特征融合, 进一步优化遮挡约束特征图; 最后将优化后的特征图输入残差流解码器计算遮挡优化后的残差流, 并将残差流与上采样后的上一级光流融合获取最终的遮挡优化光流. 由于变形特征图经遮挡优化后, 可以减少匹配代价卷中的误匹配信息, 因此能够有效抑制图像遮挡区域的光流计算结果中的运动边缘模糊现象.

### 2.3 基于特征变形的光流残差补偿模块

针对现有的深度学习光流计算网络模型在大位移场景下光流计算准确性较低的问题, 本文提出一种基于特征变形的光流残差补偿残模块, 该模块通过将残差流与遮挡优化光流融合, 实现了大位移运动区域光流信息补偿, 进而提升大位移运动光流计算的准确性与可靠性.

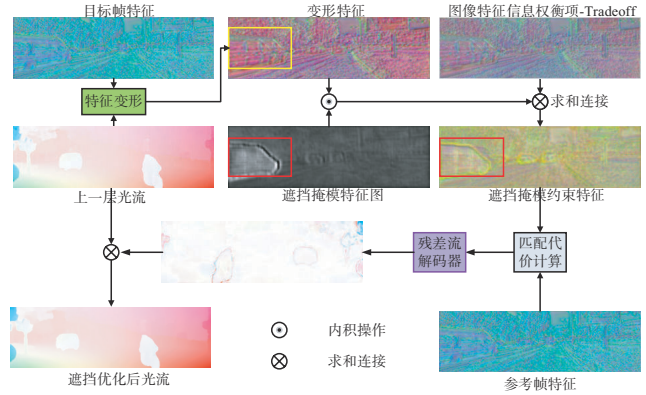


图3 基于遮挡掩模的光流约束优化过程

图4展示了基于特征变形的光流残差补偿模块网络架构示意图, 首先利用经过遮挡优化后的光流对目标帧进行特征变形, 并将计算得到的变形特征和参考帧特征、遮挡优化后的光流在特征通道上进行特征级联, 然后将级联特征输入残差流解码器计算残差光流, 最后将残差光流与遮挡优化后的光流进行融合, 以获取细化后的计算光流. 图4中残差流解码器的网络结构与图3中遮挡解码器的网络结构相似, 都由四个卷积模块和一个卷积层组成, 不同之处在于, 遮挡解码器中的卷积模块输入与输出均需要通过特征级联, 实现浅层特征信息与深层特征信息融合, 从而促进遮挡掩模光流约束模块学习到两帧图像间更准确的运动遮挡信息; 而残差流解码器中每个卷积模块的输出直接传递到下一个卷积模块进行特征提取, 这种方式相对前者既避免了重复特征的冗余融合也有效地降低模型参数量.

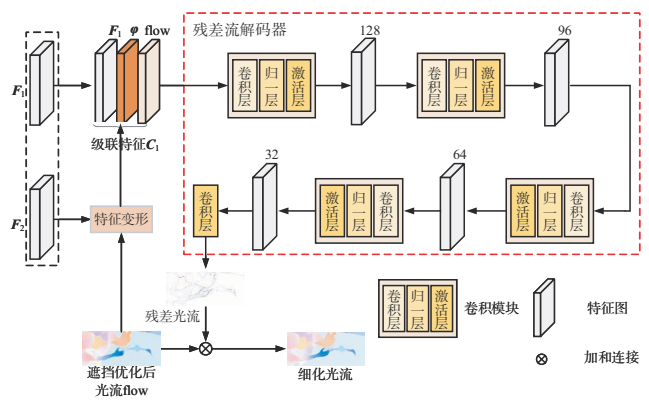


图4 基于特征变形的光流残差补偿模块

## 3 模型损失函数与训练策略

### 3.1 模型损失函数

本文采用端点误差作为网络模型训练的损失函数, 通过计算金字塔中每一层的光流计算值与光流真实值之间的 L2 范数距离<sup>[14]</sup>作为模型训练收敛的判断

条件,具体计算公式如下所示:

$$\text{loss} = \sum_{i=1}^5 \alpha_i \sum_{x \in \Omega} |\text{flow}_i(x) - \text{flow}_{\text{GT}}(x)|_2 \quad (2)$$

式(2)中,loss表示模型训练损失函数, $\text{flow}_i(x)$ 和 $\text{flow}_{\text{GT}}(x)$ 分别表示像素点 $x$ 处的光流计算值与光流真实值, $|\cdot|_2$ 表示计算光流与真实光流的L2范数距离,符号 $\Omega$ 表示整个图像区域.符号 $\alpha_i$ 表示金字塔第 $i$ 层光流计算损失函数权重系数,训练中本文分别设置为 $\alpha_1 = 1.0, \alpha_2 = 0.5, \alpha_3 = 0.25, \alpha_4 = 0.25, \alpha_5 = 0.25$ <sup>[16]</sup>.

### 3.2 模型训练策略

本文采用Adam<sup>[26]</sup>优化器最小化模型训练损失函数.首先,为了使模型学习旋转、平移以及缩放等刚性运动的二维运动特征,本文将模型在具有旋转、平移、缩放图像的FlyingChairs<sup>[13]</sup>数据集上进行训练;然后,为了使模型学习到丰富的三维空间特征,本文将模型在具有复杂三维运动的FlyingThings3D<sup>[27]</sup>数据集上进行微调训练.经过两个阶段的训练,模型可以学习到丰富的二维运动特征和三维运动特征,显著增强模型的鲁棒性;最后,在预训练模型的基础上,针对MPI-Sintel<sup>[28]</sup>和KITTI数据集上进行相应的微调训练,提高模型的泛化性,避免模型过拟合.

## 4 实验结果与分析

### 4.1 光流评价指标

本文分别采用平均端点误差(Average End-point Error, AEE)和异常值百分比(Percentage of Flow Error, Fl-all)对MPI-Sintel和KITTI数据集的光流计算结果进行量化对比与分析.其中,平均端点误差表示光流计算值与真实值的平均偏移距离,异常值百分比表示光流计算值偏离真实值超过设定阈值的离群像素点占整幅图像的百分比,计算公式如下:

$$\text{AEE} = \frac{1}{N} \sqrt{(\mathbf{u}_{\text{GT}} - \mathbf{u}_{\text{E}})^2 + (\mathbf{v}_{\text{GT}} - \mathbf{v}_{\text{E}})^2} \quad (3)$$

$$\text{Fl-all} = \frac{\sum P(\text{AEE} > \tau)}{\text{ALL}} \times 100\% \quad (4)$$

式(3)中,AEE表示光流平均端点误差, $(\mathbf{u}_{\text{E}}, \mathbf{v}_{\text{E}})^T$ 表示光流计算结果, $(\mathbf{u}_{\text{GT}}, \mathbf{v}_{\text{GT}})^T$ 为光流真实值, $N$ 表示图像中像素点个数.式(4)中,Fl-all表示光流异常值百分比, $P(\text{AEE} > \tau)$ 表示光流端点误差大于 $\tau$ 的像素点,其中 $\tau = 3$ .ALL表示图像所有的像素个数.

### 4.2 对比方法

为验证本文方法光流计算的准确性与鲁棒性,分别选取当前具有代表性的传统光流计算方法JOF<sup>[7]</sup>(Optical Flow estimation using Joint filtering)、CPM-Flow<sup>[29]</sup>(Coarse-to-fine Patch Match for optical Flow)、EpicFlow<sup>[30]</sup>(Edge-preserving interpolation of correspon-

dences for optical Flow)以及深度学习光流计算方法FlowNet2<sup>[14]</sup>(optical Flow estimation Network 2)、PWC-Net<sup>[16]</sup>(Pyramid, Warping, and Cost volume Network)、LiteFlowNet2<sup>[17]</sup>(Lightweight Convolutional Neural Network 2 for optical Flow estimation)、VCN<sup>[20]</sup>(Volumetric Correspondence Networks)、UnFlow<sup>[22]</sup>(Unsupervised learning of optical Flow)、DDFlow<sup>[23]</sup>(Data Distillation Flow)与本文方法进行综合对比分析.其中,JOF设计基于多滤波协同优化策略的变分光流计算模型,可以实现较好的光流边缘保护效果.CPM-Flow提出了一种由粗到细的块匹配光流计算方法,通过融合有效随机搜索策略和由粗到细的光流计算方案,提高了大位移场景光流计算的精度.EpicFlow通过从初始稀疏匹配集执行稀疏到密集的插值,利用边缘感知测地线距离的轮廓线索来计算稠密光流,在提高大位移与遮挡场景光流计算精度的同时保护了光流结果的边缘轮廓信息.FlowNet2方法将多个不同功能的子网络堆叠实现优化大位移运动和小位移运动区域的光流计算.PWC-Net方法在特征金字塔后置一个上下文光流补偿网络,有效提升了模型的光流计算精度.LiteFlowNet2方法提出一个轻型描述子匹配光流和特征驱动正则化光流相结合的网络模型,有效抑制了边缘模糊现象.UnFlow无监督方法将遮挡信息作为损失函数的附加约束项,对遮挡场景具有较好的鲁棒性.DDFlow方法利用数据驱动的蒸馏学习方式对模型进行训练,有效提高了无监督网络模型的光流计算精度.VCN方法采用可分离四维卷积滤波处理的匹配代价关系,在减少计算量的同时提高了光流计算模型的精度.

### 4.3 MPI-Sintel数据集实验对比

为验证本文方法针对运动遮挡和大位移等复杂场景的光流计算精度与鲁棒性,本文首先采用由计算机合成的MPI-Sintel数据集对本文方法和对比方法进行实验测试与对比分析.表1列出了本文方法和各对比方法在MPI-Sintel数据集的光流计算误差及时间消耗对比,其中,all表示整幅图像的光流平均端点误差,matched和unmatched分别表示非遮挡和遮挡区域的光流平均端点误差.从表1中可以看出,本文方法在整幅图像的光流平均端点误差仅略低于VCN方法,说明本文方法具有较高的光流计算精度.值得注意的是,本文方法在图像遮挡区域的平均端点误差最小,说明本文通过在解码网络中引入基于遮挡掩模的光流约束模块可以有效提升遮挡区域光流计算的精度.其次,在时间消耗方面,传统光流计算方法相对深度学习光流计算方法在时耗方面具有明显劣势,深度学习方法中PWC-Net方法的时间消耗最少,但该方法的光流计算整体精度低于本文方法与LiteFlowNet2.本文方法由于采用遮

挡约束和残差补偿对光流计算进行深度优化,因此时间消耗略大于其他对比方法。

表2分别列出了本文方法与各对比方法在不同遮挡边界与大位移区域的光流计算结果。其中,d0-10、d10-60和d60-140分别表示距离遮挡边界不同像素距离区域内光流计算的平均端点误差,s0-10,s10-40和s40+分别表示图像连续帧间不同位移距离区域的光流计算平均端点误差。其中,鉴于d0-10指示了光流计算在遮挡边界0~10像素区域的平均端点误差,s10-40和s40+分别指示了光流计算在位移大于10个像素和40个像素以上区域的平均端点误差。因此,d0-10、s10-40和s40+指标直接反映了光流计算在运动遮挡和大位移区域的表现。从表2中可以看出,本文方法在d0-10和s10-40指标精度最高,在s40+误差与VCN方法相近。由于这些指标中像素遮挡区域较大且像素点位移相对较为明显,因此,本文提出的遮挡掩模光流约束模块生成的遮挡权重项对该区域像素点作用更为显著,变形特征矫正也更加精确,进而也促进了特征变形的光流残差补偿模块对大位移运动光流信息的补偿。

图5展示了本文方法和JOF、CPM-Flow、EpicFlow、PWC-Net、LiteFlowNet2、VCN等对比方法针对MPI-Sintel数据集Perturbed Shaman\_1、Ambush\_3序列的光流计算可视化结果。图中黄色方框标签区域为运动遮挡与大位移区域,图下方为对应标签区域局部放大图。从图5(a)中可以看出,JOF、PWC-Net以及VCN方法在运动遮挡区域存在明显的错误计算,CPM-Flow、EpicFlow和LiteFlowNet2方法虽然较好的计算出运动遮挡区域的轮廓信息,但边缘模糊现象较为严重。本文方法相对其他方法更加精准的计算出运动遮挡区域的光流信息,边缘模糊现象得到有效改善。从图5(b)可以看出,传统光流计算方法在标签区域均存在较大范围的错误计算且异常值较多,深度学习方法PWC-Net、LiteFlowNet2和VCN虽然相对传统方法较好的计算出大位

移运动区域光流,但是边缘模糊现象仍然较为明显。本文方法通过引入遮挡掩模的光流约束模块与特征变形的光流残差补偿模块不仅更高精度的预测出该区域完整光流信息,而且有效抑制了边缘模糊现象。

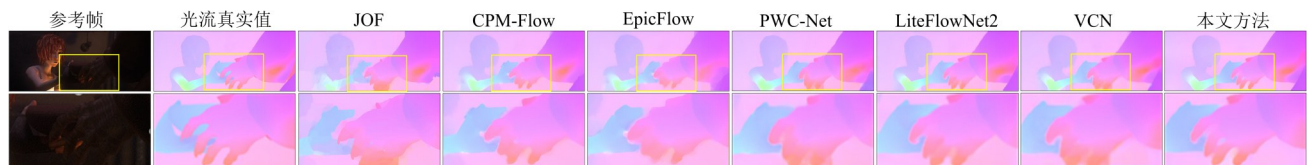
表1 MPI-Sintel数据集光流计算结果及时间消耗对比

对比方法	all /pixel	matched /pixel	Unmatched /pixel	时间 /s
JOF <sup>[17]</sup>	8.818	4.599	43.175	654
CPM-Flow <sup>[29]</sup>	5.960	2.990	30.177	4.3
EpicFlow <sup>[30]</sup>	6.285	3.060	32.564	16
FlowNet2 <sup>[14]</sup>	6.016	2.977	30.807	0.10
PWC-Net <sup>[16]</sup>	5.042	2.445	26.221	<b>0.03</b>
LiteFlowNet2 <sup>[17]</sup>	4.686	2.248	24.571	0.04
UnFlow <sup>[22]</sup>	10.219	6.061	44.110	0.12
DDFlow <sup>[23]</sup>	6.176	2.269	38.053	0.06
VCN <sup>[20]</sup>	<b>4.404</b>	<b>2.216</b>	22.238	0.18
本文方法	4.607	2.482	<b>21.923</b>	0.20

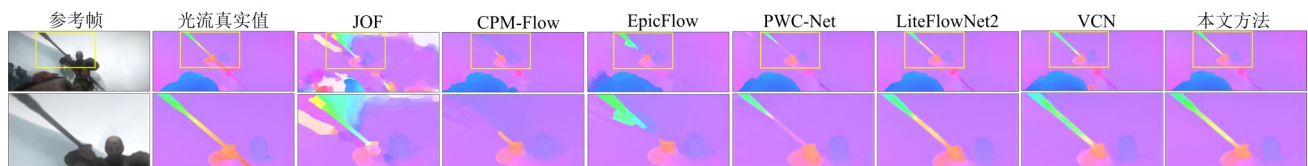
#### 4.4 KITTI数据集实验对比

为进一步验证本文方法在真实场景下的光流计算准确性与鲁棒性,采用KITTI数据集对本文方法和对比方法进行实验测试与对比分析。KITTI数据集<sup>[31]</sup>由静态背景的KITTI2012数据集和动态背景的KITTI2015数据集组成。相比KITTI2012数据集,KITTI2015数据集包含更多的光照变化、大位移、运动遮挡以及运动模糊等场景,能够更好地反映不同计算方法在真实场景下的光流计算准确性与鲁棒性。

表3分别列出了本文方法与各对比方法在该数据集的光流计算结果对比。表中KITTI2015数据集提供了图像背景区域(bg)、图像前景区域(fg)和图像全部区域(all)的光流异常值百分比等量化指标,KITTI2012数据集提供了图像非遮挡区域(noc)和图像全部区域(all)的光流异常值百分比等量化指标。从表3可以看出,传统方法CPM-Flow与EpicFlow方法在KITTI2015



(a) Perturbed Shaman\_1运动遮挡图像序列光流估计结果及其细节放大图对比



(b) Ambush\_3运动遮挡图像序列光流估计结果及其细节放大图对比

图5 MPI-Sintel数据集图像序列光流计算可视化结果对比

表 2 MPI-Sintel 数据集遮挡边界与大位移区域光流计算结果

对比方法	d0-10/pixel	d10-60/pixel	d60-140/pixel	s0-10/ppf	s10-40/ppf	s40+/ppf
JOF <sup>[17]</sup>	7.049	4.617	3.131	1.170	4.196	57.923
CPM-Flow <sup>[29]</sup>	5.038	2.419	2.143	1.155	3.755	35.136
EpicFlow <sup>[30]</sup>	5.205	2.611	2.216	1.135	3.727	38.021
FlowNet2 <sup>[14]</sup>	5.139	2.786	2.102	1.243	4.027	34.505
PWC-Net <sup>[16]</sup>	4.636	2.087	1.475	<b>0.799</b>	2.986	31.070
LiteFlowNet2 <sup>[17]</sup>	4.048	1.899	1.473	0.811	2.433	29.375
UnFlow <sup>[22]</sup>	8.407	5.828	4.665	1.742	6.689	60.765
DDFlow <sup>[23]</sup>	4.208	2.084	<b>1.416</b>	0.860	2.562	41.337
VCN <sup>[20]</sup>	4.381	<b>1.782</b>	1.423	0.955	2.725	<b>25.570</b>
本文方法	<b>3.935</b>	1.935	1.842	1.082	<b>2.387</b>	27.462

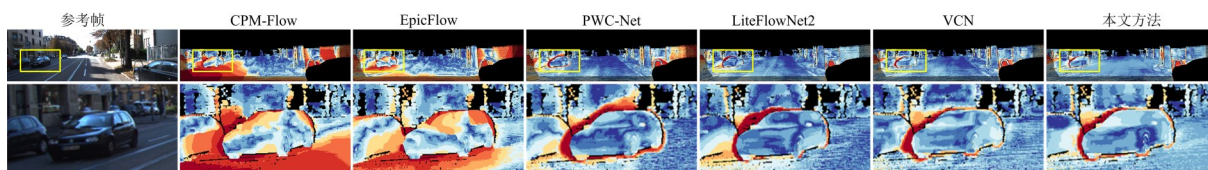
数据集整体测试精度低于深度学习方法,而在 KITTI2012 数据集传统方法在 Fl-all 指标优于 FlowNet2 与 PWC-Net,但在 Fl-noc 指标整体处于劣势.这说明传统光流算法针对简单运动真实场景具有较好效果,但是随着场景复杂度的增加光流计算精度急剧降低.本文方法在 KITTI2012 和 KITTI2015 数据集除 Fl-fg 指标外均取得了最优表现,原因是 KITTI 数据集包含了更多阴影区域而阴影区域与伪影存在一定相似,因此导致本文方法在 Fl-fg 指标低于 LiteFlowNet2 方法.其次,虽然 VCN 方法在 MPI-Sintel 数据集的 matched 指标略优于本文方法,但本文方法在真实场景 KITTI 数据集表现明显优于 VCN 方法与其他对比方法,说明本文方法具有较高的光流计算精度与鲁棒性.

图 6 展示了本文方法与对比方法在 KITTI2015 数据集 000004 和 000009 序列的光流计算误差可视化对比,图中颜色越接近红色表示光流误差越大,颜色越接近蓝色表示光流误差越小.从图中可以看出,传统光流算法 CPM-Flow、EpicFlow 在运动遮挡区域和大位移运动区域整体光流计算误差最大.由于 PWC-Net 模型未针对遮挡场景进行优化,因此在图像序列中的前景车

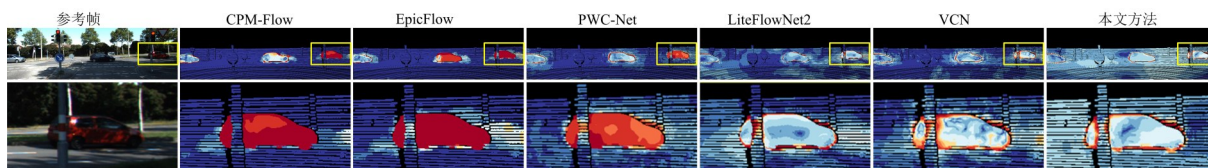
辆运动边缘计算效果差,尤其是 000009 序列中阴影遮挡处的车辆明显计算错误. LiteFlowNet2 模型采取的特征驱动局部卷积策略只是将局部邻域的光流信息扩散到计算错误的区域,前景运动车辆的边缘和大位移区域的光流计算效果依然有待改善. VCN 模型仅对全局匹配代价信息进行滤波处理,并不能有效改善运动遮挡和大位移场景下的光流计算效果.本文方法通过设计专门应对运动遮挡与大位移运动的遮挡掩模的光流约束模块与特征变形的光流残差补偿模块,使得在这两个区域本文方法光流计算误差相比于其他对比方法最小.根据 KITTI 数据集的光流计算结果定性对比可知,本文方法能够显著提高运动遮挡和大位移场景的光流计算精度与鲁棒性.

#### 4.5 消融实验

为验证本文方法提出的基于遮挡掩模的光流约束模块和基于特征变形的光流残差补偿模块对光流计算精度与鲁棒性的提升作用,采用 KITTI2015 数据集对本文方法进行消融实验对比分析.表 4 列出了本文方法和不同消融模型在 KITTI2015 测试数据集的光流计算误差对比结果,其中, Baseline 表示基于可分离四维卷



(a) 000004 大位移图像序列光流估计结果及其细节对比放大图



(b) 000009 运动遮挡图像序列光流估计结果及其细节放大图对比

图 6 KITTI2015 数据集图像序列光流计算误差可视化结果对比

表3 KITTI数据集图像序列光流计算结果对比 单位:%

对比方法	KITTI2015			KITTI2012	
	Fl-all	Fl-fg	Fl-bg	Fl-all	Fl-noc
CPM-Flow <sup>[29]</sup>	22.32	22.81	22.40	5.79	13.70
EpicFlow <sup>[30]</sup>	25.81	28.69	26.29	7.88	17.08
FlowNet2 <sup>[14]</sup>	10.41	8.75	10.75	8.80	4.82
PWC-Net <sup>[16]</sup>	9.60	9.31	9.66	8.10	4.22
LiteFlowNet2 <sup>[17]</sup>	7.62	<b>7.64</b>	7.62	6.16	2.63
UnFlow <sup>[22]</sup>	11.11	15.93	10.15	8.42	4.28
DDFlow <sup>[23]</sup>	14.29	20.40	13.08	7.34	4.57
VCN <sup>[20]</sup>	6.30	8.66	5.83	5.64	2.48
本文方法	<b>5.91</b>	8.11	<b>5.46</b>	<b>5.54</b>	<b>2.38</b>

积滤波的光流计算基准模型<sup>[21]</sup>, Base\_Occ 表示基准模型耦合遮挡掩模的光流约束模块后的网络模型, Base\_ResF 表示基准模型特征变形的光流残差补偿模块后的网络模型, 本文方法是基准模型耦合遮挡掩模的光流约束模块与特征变形的光流残差补偿模块后的全模型。

由表4可以看出, Base\_Occ 模型在图像整体区域和背景区域的光流异常值百分比略高于 Baseline 模型, 这是因为遮挡掩模权重的误差会引起非遮挡区域光流产生错误。但是, Base\_Occ 模型在图像前景区域的异常值百分比相对 Baseline 模型有所降低, 鉴于运动遮挡一般存在于图像前景目标区域, 说明遮挡约束模块能够有效提升图像前景目标遮挡区域的光流计算精度。相较于 Baseline 模型, Base\_ResF 模型在图像整体区域和背景区域的光流异常值百分比明显降低, 说明残差补偿模块能够有效提升图像大面积区域的光流计算精度。在联合遮挡约束和残差补偿模块后, 本文方法在图像整体区域、前景区域和背景区域的光流计算精度显著高于 Baseline 模型, 说明本文提出的基于可学习遮挡掩模的特征变形约束模块和基于特征变形的光流残差补偿模块能够有效提高光流计算的整体精度, 尤其在运动遮挡和大位移区域具有更好的准确性与鲁棒性。此外, 在加入遮挡约束和残差补偿模块后, 本文方法及不同消融模型的训练和测试时间并未明显增加, 说明本文提出的遮挡约束和残差补偿模块未显著增大模型的训练和计算复杂度。

表4 本文方法不同消融模型光流计算对比结果

对比模型	Fl-bg /%	Fl-fg /%	Fl-all /%	训练 时间/h	测试 时间/s
Baseline	5.83	8.66	6.30	227	0.18
Base_Occ	6.13	8.60	6.54	255	0.19
Base_ResF	<b>5.44</b>	8.17	<b>5.89</b>	241	0.19
本文方法	5.46	<b>8.11</b>	5.91	279	0.20

图7展示了本文方法和不同消融模型针对000007图像序列的光流计算误差对比结果。从图中可以看出, Base\_Occ 与 Base\_ResF 消融模型能够较好地优化前景运动车辆和背景区域的光流计算效果, 本文方法在联合遮挡约束和残差补偿两个模块后, 能够显著提升图中运动遮挡和大位移区域的光流计算效果。

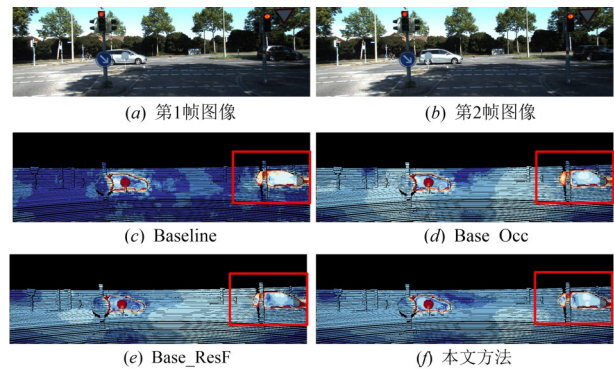


图7 不同消融模型光流误差对比结果

## 5 结论

本文提出一种联合遮挡约束与残差补偿的特征金字塔光流计算方法。首先, 设计基于遮挡掩模的光流约束模块, 利用该模块输出的遮挡掩模抑制变形特征的边缘伪影, 克服运动遮挡区域的图像边缘模糊问题。然后, 基于特征变形策略设计光流残差补偿模块, 利用该模块计算的残差流细化原始光流, 改善大位移运动区域的光流计算效果。最后, 采用特征金字塔框架构建耦合遮挡约束与残差补偿的深度学习光流计算网络模型。分别采用 MPI-Sintel 和 KITTI 数据集对本文方法与现有的代表性深度学习光流计算方法进行综合对比分析, 实验结果表明本文方法针对运动遮挡和大位移场景具有更好的光流计算精度与鲁棒性。

## 参考文献

- [1] 吴晓军, 鞠光亮. 一种无标记点人脸表情捕捉与重现算法[J]. 电子学报, 2016, 44(9): 2141-2147.  
WU X J, JU G L. A markerless facial expression capture and reproduce algorithm[J]. Acta Electronica Sinica, 2016, 44(9): 2141-2147. (in Chinese)
- [2] JI Y L, YANG Y, SHEN F M, et al. Arbitrary-view human action recognition: A varying-view RGB-D action dataset [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(1): 289-300.
- [3] 丁新尧, 张鑫. 基于显著性特征的选择性目标跟踪算法 [J]. 电子学报, 2020, 48(1): 118-123.  
DING X Y, ZHANG X. Visual tracking with salient features and selective mechanism[J]. Acta Electronica Sinica,

- 2020, 48(1): 118-123. (in Chinese)
- [4] 张聪炫, 陈震, 黎明. 单目图像序列光流三维重建技术研究综述[J]. 电子学报, 2016, 44(12): 3044-3052.  
ZHANG C X, CHEN Z, LI M. Review of the 3D reconstruction technology based on optical flow of monocular image sequence[J]. Acta Electronica Sinica, 2016, 44(12): 3044-3052. (in Chinese)
- [5] HORN B K P, SCHUNCK B G. Determining optical flow [J]. Artificial Intelligence, 1981, 17(1/2/3): 185-203.
- [6] MEI L, LAI J H, XIE X H, et al. Illumination-invariance optical flow estimation using weighted regularization transform[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30(2): 495-508.
- [7] ZHANG C X, GE L Y, CHEN Z, et al. Refined TV-L1 optical flow estimation using joint filtering[J]. IEEE Transactions on Multimedia, 2020, 22(2): 349-364.
- [8] SUN D Q, ROTH S, BLACK M J. A quantitative analysis of current practices in optical flow estimation and the principles behind them[J]. International Journal of Computer Vision, 2014, 106(2): 115-137.
- [9] ZHANG C X, CHEN Z, WANG M R, et al. Robust non-local TV-L1 optical flow estimation with occlusion detection [J]. IEEE Transactions on Image Processing, 2017, 26(8): 4055-4067.
- [10] ZIMMER H, BRUHN A, WEICKERT J. Optic flow in harmony[J]. International Journal of Computer Vision, 2011, 93(3): 368-388.
- [11] SEVILLA-LARA L, SUN D Q, JAMPANI V, et al. Optical flow with semantic segmentation and localized layers [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 3889-3898.
- [12] 张聪炫, 周仲凯, 陈震, 等. 深度学习光流计算技术研究进展[J]. 电子学报, 2020, 48(9): 1841-1849.  
ZHANG C X, ZHOU Z K, CHEN Z, et al. Research progress of deep learning based optical flow computation technology[J]. Acta Electronica Sinica, 2020, 48(9): 1841-1849. (in Chinese)
- [13] DOSOVITSKIY A, FISCHER P, ILG E, et al. FlowNet: learning optical flow with convolutional networks[C]//2015 IEEE International Conference on Computer Vision. Boston: IEEE, 2015: 2758-2766.
- [14] ILG E, MAYER N, SAIKIA T, et al. FlowNet 2.0: evolution of optical flow estimation with deep networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Hawaii: IEEE, 2017: 1647-1655.
- [15] RANJAN A, BLACK M J. Optical flow estimation using a spatial pyramid network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 2720-2729.
- [16] SUN D Q, YANG X D, LIU M Y, et al. PWC-net: CNNs for optical flow using pyramid, warping, and cost volume [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 8934-8943.
- [17] HUI T W, TANG X O, LOY C C. LiteFlowNet: A lightweight convolutional neural network for optical flow estimation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 8981-8989.
- [18] ZHAO S Y, SHENG Y L, DONG Y, et al. MaskFlowNet: asymmetric feature matching with learnable occlusion mask[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 6277-6286.
- [19] LU Y, VALMADRE J, WANG H, et al. Devon: Deformable volume network for learning optical flow[C]//2020 IEEE Winter Conference on Applications of Computer Vision. Snowmass: IEEE Press, 2020. 2705-2713.
- [20] YANG G, RAMANAN D. Volumetric correspondence networks for optical flow[C]//33rd International Conference on Neural Information Processing Systems. Vancouver: Curran Associates Inc, 2019: 794-805.
- [21] YU J J, HARLEY A W, DERPANIS K G. Back to basics: Unsupervised learning of optical flow via brightness constancy and motion smoothness[C]//European Conference on Computer Vision. Cham: Springer, 2016: 3-10.
- [22] MEISTER S, HUR J, ROTH S. UnFlow: Unsupervised learning of optical flow with a bidirectional census loss [C]//AAAI Conference on Artificial Intelligence. New Orleans: AAAI Press, 2018: 7251-7259.
- [23] LIU P P, KING I, LYU M R, et al. DDFlow: learning optical flow with unlabeled data distillation[J]. Proceedings of the AAAI Conference on Artificial Intelligence. Hawaii: AAAI Press, 2019, 33: 8770-8777.
- [24] LAI W S, HUANG J B, YANG M H. Semi-supervised learning for optical flow with Generative Adversarial Networks[C]//31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc, 2017: 353-363.
- [25] SONG X L, ZHAO Y Y, YANG J Y, et al. FPCR-net: Feature pyramidal correlation and residual reconstruction for optical flow estimation[EB/OL]. (2020-01-17). <https://>

arxiv.org/abs/2001.06171.

- [26] KINGMA D, BA J. Adam: A method for stochastic optimization[C]//International Conference for Learning Representations. San Diego: Elsevier Press, 2015: 1-15.
- [27] MAYER N, ILG E, HÄUSSER P, et al. A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 4040-4048.
- [28] BUTLER D J, WULFF J, STANLEY G B, et al. A naturalistic open source movie for optical flow evaluation[C]//European Conference on Computer Vision. Berlin, Heidelberg: Springer, 2012: 611-625.
- [29] HU Y L, SONG R, LI Y S. Efficient coarse-to-fine patch match for large displacement optical flow[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 5704-5712.
- [30] REVAUD J, WEINZAEPFEL P, HARCHAOUI Z, et al. EpicFlow: Edge-preserving interpolation of correspondences for optical flow[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 1164-1172.
- [31] MENZE M, GEIGER A. Object scene flow for autonomous vehicles[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 3061-3070.



张聪炫(通讯作者) 男,1984年7月出生  
于河南省焦作市. 2014年在南京航空航天大学  
获博士学位. 现为南昌航空大学教授,硕士生导  
师. 主要研究方向为图像检测与智能识别. 中  
国电子学会会员编号:E190014621M.

E-mail: zcxdsq@163.com



陈震 男,1969年11月出生  
于江西省九江市. 2003年在西北工业大学获得博士学位.  
现为南昌航空大学教授,博士生导师. 主要研  
究方向为计算机视觉、图像处理与模式识别.

E-mail: dr\_chenzhen@163.com



黎明 男,1965年2月出生  
于江西省樟树市. 1997年获南京航空航天大学博士学位.  
现为南昌航空大学教授,博士生导师. 主要研  
究方向为图像处理与模式识别、智能计算.

E-mail: limingniat@hotmail.com

#### 作者简介



范兵兵 男,1997年3月出生  
于江西省瑞昌市. 现为南昌航空大学测试与光电工程学院  
硕士研究生. 主要研究方向为图像检测与智能  
识别.

E-mail: 2652459616@qq.com



何庭建 男,1995年4月出生  
于江西省抚州市, 2021年6月南昌航空大学信息工程学院  
硕士研究生毕业,主要研究方向为图像检测与  
智能识别.

E-mail: 511076738@qq.com