

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

联合空角信息的无参考光场图像质量评价

王斌, 白永强, 朱仲杰, 郁梅, 蒋刚毅

引用本文:

王斌, 白永强, 朱仲杰, 等. 联合空角信息的无参考光场图像质量评价[J]. 光电工程, 2024, 51(9): 240139.

Wang B, Bai Y Q, Zhu Z J, et al. No-reference light field image quality assessment based on joint spatial-angular information[J]. *Opto-Electron Eng*, 2024, 51(9): 240139.

<https://doi.org/10.12086/oe.2024.240139>

收稿日期: 2024-06-14; 修改日期: 2024-08-18; 录用日期: 2024-08-18

相关论文

LF-UMTI: 基于多尺度空角交互的无监督多曝光光场图像融合

李玉龙, 陈晔曜, 崔跃利, 郁梅

光电工程 2024, 51(6): 240093 doi: [10.12086/oe.2024.240093](https://doi.org/10.12086/oe.2024.240093)

角度差异强化的光场图像超分网络

吕天琪, 武迎春, 赵贤凌

光电工程 2023, 50(2): 220185 doi: [10.12086/oe.2023.220185](https://doi.org/10.12086/oe.2023.220185)

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](https://doi.org/10.12086/oe.2021.210405)

融合空频复用和近远场复用的多功能超表面图像显示

周宇翀, 丁玮珺, 李子乐, 刘宏超, 付娆, 戴琦, 郑国兴

光电工程 2023, 50(8): 230153 doi: [10.12086/oe.2023.230153](https://doi.org/10.12086/oe.2023.230153)

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

联合空角信息的无参考光场图像质量评价

王 斌¹, 白永强², 朱仲杰², 郁 梅^{1*}, 蒋刚毅¹

¹宁波大学 信息科学与工程学院, 浙江 宁波 315211;

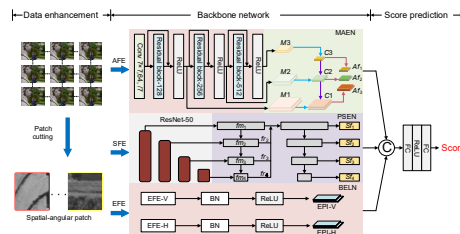
²浙江万里学院 信息与智能工程学院, 浙江 宁波 315100

摘要: 光场图像通过记录多个视点信息可为用户提供更加全面真实的视觉体验, 但采集和可视化过程中引入的失真会严重影响其视觉质量。因此, 如何有效地评价光场图像质量是一个巨大挑战。本文结合空间-角度特征和极平面信息提出了一种基于深度学习的无参考光场图像质量评价方法。首先, 构建了空间-角度特征提取网络, 通过多级连接以达到捕获多尺度语义信息的目的, 并采用多尺度融合方式实现双重特征有效提取; 其次, 提出双向极平面图像特征学习网络, 以有效评估光场图像角度一致性; 最后, 通过跨特征融合并线性回归输出图像质量分数。在三个通用数据集上的对比实验结果表明, 所提出方法明显优于经典的 2D 图像和光场图像质量评价方法, 其评价结果与主观评价结果的一致性更高。

关键词: 光场图像; 空间-角度特征; 极平面信息; 无参考图像质量评价; 角度一致性

中图分类号: TP394.1

文献标志码: A



王斌, 白永强, 朱仲杰, 等. 联合空角信息的无参考光场图像质量评价[J]. 光电工程, 2024, 51(9): 240139

Wang B, Bai Y Q, Zhu Z J, et al. No-reference light field image quality assessment based on joint spatial-angular information[J]. *Opto-Electron Eng*, 2024, 51(9): 240139

No-reference light field image quality assessment based on joint spatial-angular information

Wang Bin¹, Bai Yongqiang², Zhu Zhongjie², Yu Mei^{1*}, Jiang Gangyi¹

¹Faculty of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China;

²College of Information and Intelligent Engineering, Zhejiang Wanli University, Ningbo, Zhejiang 315100, China

Abstract: Light field images provide users with a more comprehensive and realistic visual experience by recording information from multiple viewpoints. However, distortions introduced during the acquisition and visualization process can severely impact their visual quality. Therefore, effectively evaluating the quality of light field images is a significant challenge. This paper proposes a no-reference light field image quality assessment method based on deep learning, combining spatial-angular features and epipolar plane information. Firstly, a spatial-angular feature extraction network is constructed to capture multi-scale semantic information through multi-level connections, and a multi-scale fusion approach is employed to achieve effective dual-feature extraction. Secondly, a bidirectional epipolar plane image feature learning network is proposed to effectively assess the angular consistency of light field images. Finally, image quality scores are output through cross-feature fusion and linear regression. Comparative

收稿日期: 2024-06-14; 修回日期: 2024-08-18; 录用日期: 2024-08-18

基金项目: 国家自然科学基金资助项目 (62271276)

*通信作者: 郁梅, yumei@nbu.edu.cn。

版权所有©2024 中国科学院光电技术研究所

experimental results on three common datasets indicate that the proposed method significantly outperforms classical 2D image and light field image quality assessment methods, with a higher consistency with subjective evaluation results.

Keywords: light field images; spatial-angular feature; epipolar plane information; no-reference image quality assessment; angular consistency

1 引言

伴随着沉浸式多媒体技术的飞速发展, 光场成像已成为计算视觉领域的研究热点^[1]。与传统 2D 图像不同, 光场图像 (light field image, LFI) 不仅可以呈现出图像的空间域信息, 同时还包括独特的角度域信息。光场图像信号可以由 7D 全光函数数学模型表示, 其可进一步简化为双平面光场模型 $\mathcal{L} \in \mathbb{R}^{u \times v \times x \times y}$, 即用 4D 函数表示, 其中 (u, v) 平面和 (x, y) 平面分别对应于角度域信息和空间域信息^[2]。因为其丰富的空间和角度信息, 光场广泛应用于图像应用, 包括深度估计、重聚焦、3D 重建等^[3-4]。然而在光场图像处理过程中, 压缩、重建等操作都会引入各种复杂失真, 从而降低用户视觉感知体验质量。因此, 研究准确的光场图像质量评价 (light field image quality assessment, LF-IQA) 方法以优化光场图像处理方法与过程是十分必要的。

光场图像质量评价根据原始参考图像信息的参与程度可以分为三类。其中全参考 (full reference, FR) 和半参考 (reduced reference, RR) LF-IQA 使用全部或部分原始参考信息用来评估失真 LFI 的视觉感知质量。例如, Tian 等人^[5]提出了一种基于多阶导数特征模型 (multi-order derivative decision-based model, MDFM), 计算原始光场图像与失真光场图像中的所有子孔径图像 (sub-aperture image, SAI) 导数特征的相似度平均值作为失真光场图像的分值。Huang 等人^[6]基于人眼对空间视角下的尺度、方向、轮廓和几何视角下的视差变化更感兴趣的事实, 使用轮廓波变换在参考和失真 SAI 上提取空间特征, 同时使用 3D Gabor 滤波器在参考和失真 SAI 序列提取角度特征对参考 LFI 和失真 LFI 进行相似度度量。Min 等人^[7]从全局空间质量、局部空间质量以及角度质量三个方面量化失真光场图像质量, 利用 SAI 的结构匹配度和近边缘区域的均方误差评估光场图像的全局和局部空间质量, 结合视图密度和视图质量变化信息来估计光场图像的角度质量。然而在某些实际应用中, 参考 LFI 可能无法获得, 使得全参考和半参考方法的使用受到限制。而无参考

(blind/no reference, NR) LF-IQA 则是不需要原始参考信息的参与直接对失真 LFI 进行视觉质量评估, 因而具有更广泛的应用场景。

受到 2D 图像质量评价方法的启发, 前期有关 NR LF-IQA 的工作更多的是采用在 LFI 的 2D 表示中提取传统手工特征的方式, 分别获取空间/角度的失真信息, 再通过支持向量回归或随机森林方式训练模型来评估 LFI 的感知质量分数。例如, Shi 等人^[8]通过捕捉光场眼合成图像阵列的自然分布和利用光场极平面图像 (epipolar plane image, EPI) 的梯度方向图分布分别捕获空间和角度方向的失真信息, 衡量失真 LFI 的感知质量。Luo 等人^[9]利用全局熵分布和均匀局部二值模式描述符从微透镜图像 (micro lens image, MLI) 中提取特征, 通过信息熵计算评估 LFI 的感知质量。Xiang 等人^[10]对 LFI 进行 Curvelet 变换, 从 LFI 的均值差图像和子孔径图像中提取相应的自然场景统计特征和能量特征用于 LFI 质量预测。Lamichhane 等人^[11]通过估计 LFI 的显著性图和眼合成图来提取空间特征, 结合全局亮度分布, 在 EPI 上使用加权局部二值模式运算来获取角度特征, 实现对 LFI 的质量评估。Zhou 等人^[12]利用 Tucker 分解获得 4 个定向 SAI 堆栈的第一主成分, 提出张量角变异指数, 搭配第一主成分获取 SAI 之间的结构相似度, 综合空间和角度信息获取 LFI 质量分数。Xiang 等人^[13]基于光场图像不同的可视化方式, 在伪视频序列和重聚焦图像中提取能表征失真的特征, 并通过支持向量回归模型池化特征进行质量预测。

传统手工特征获取的方式可能会遗失掉一部分原始图像信息, 从而影响预测准确性。近些年来, 随着卷积神经网络 (convolutional neural networks, CNN) 在图像质量评价领域的逐渐应用, 很多学者提出了许多基于神经网络的 LF-IQA 方法。例如, Alamgeer 等人^[14]考虑到传统神经网络受限于局部感受野的限制, 无法长期保持学习依赖关系, 提出基于长短时记忆网络的 LF-IQA 度量方法。Qu 等人^[15]通过引入深度可

分离卷积实现对 LFI 失真程度的精准预测。为了进一步感知 LFI 的角度失真, 他们还提出了角度注意的概念和三个特定的角度注意核, 以全局或选择性地提取多角度特征, 提高了 LF-IQ 的准确性^[16]。上述方法都是基于整幅 LFI 进行特征学习, 容易面临计算量大、复杂度高的问题。为了缓解这类问题, Zhao 等^[17]使用可区分的 EPI 块来预估 LFI 质量。Zhang 等^[18]采用切块处理, 利用多个小块的图像质量分数估计整幅 LFI 的图像质量分数。

然而, 上述已知的一些研究在对空间-角度信息提取过程中没有充分考虑人眼视觉感知特性, 在对图像信息提取时, 通常采用局部或是全局特征提取, 一次性获取图像的所有信息, 并没有达到预期的性能目标。基于以上分析, 本文提出了一种联合空间-角度特征和 EPI 特征的无参考光场图像质量评价方法。首先对图像进行切块数据增强处理, 通过对图像块进行空间-角度信息分离处理, 利用两条支路分别提取空间和角度特征。在对空间信息提取时, 利用 ResNet^[19]对图像信息进行阶段化提取, 逐步实现由粗到细的特征提取。同时利用金字塔框架结构实现对角度信息的逐级特征提取, 以充分考虑局部与全局信息的关联性。为更全面地获取影响评估准确性的角度失真信息, 构建了双向 EPI 特征学习网络。在初步获取图像空间和角度信息的基础上, 增强对角度失真特征的捕捉, 提高预测精度。

2 提出的方法

2.1 光场图像的 2D 表示形式

为了更好地观察 LFI, 一般采用三种可视化方法: 微透镜图像、子孔径图像阵列、极平面图像。

微透镜图像 (MLI): 一个微透镜图像是由一系列的微透镜单元组成, 其中每一个微透镜单元聚集了所有位于空间位置 (x, y) 的光线信息, 用来获取 LFI 的角度信息, 表示为 $M_{x,y} = \{M_{x,y}(u, v)\}$, 如图 1(a) 所示。

子孔径图像阵列 (SAIs): 将 (u, v) 平面作为一组相机视图, 将 (x, y) 平面作为其焦平面, LFI 可表示为针孔视图的 2D 阵列, 并且每个视图称为 SAIs 的 2D 图像, 表示为 $I_{u,v}$, 反映了位于角坐标 (u, v) 处的空间信息, 如图 1(b) 图像阵列所示。

极平面图像 (EPI): 当同时固定 u, x 或者 v, y 时, 可以获取到 EPI 图像。EPI 图像反映了 LFI 的垂直或水平视差信息, 表示为 $E_{u,x} = \{E_{u,x}(v, y)\}$ 或 $E_{v,y} = \{E_{v,y}(u, x)\}$, 如图 1(b) 所示。

$\{E_{v,y}(u, x)\}$, 如图 1(b) 所示。

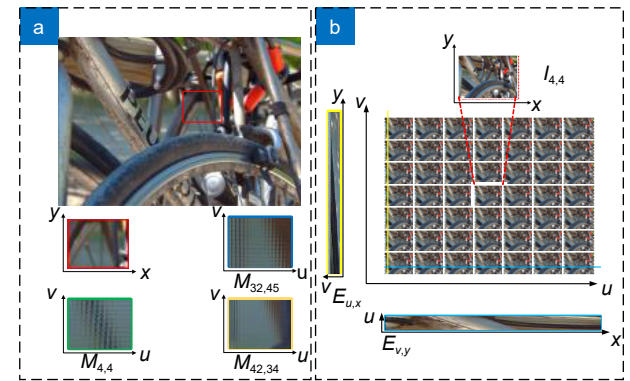


图 1 光场图像的不同表示形式。(a) MLI; (b) SAIs

Fig. 1 Different representations of light field image.

(a) MLI; (b) SAIs

图 2 为提出的联合空间-角度信息和 EPI 特征信息的无参考光场图像质量评价方法 (spatial-angular-epipolar plane information blind quality assessment of light field images, SAE-BLFI) 的总体框架及相关模块。它主要由两个部分组成: 1) 数据增强操作; 2) 多分支特征提取网络。首先, 通过数据增强操作对原始失真图像进行空间-角度图像块生成。然后, 通过渐进式空间特征提取、多尺度角度金字塔、双向 EPI 特征学习网络分别获取图像的失真特征。最后, 将三支的特征进行跨特征融合, 预测失真图形的质量分数, 再对所有的块得分进行池化操作, 得到输入失真 LFI 的总体质量得分。

2.2 数据增强

使用图像块训练深度学习模型是解决 LF-IQA 方法中训练数据不足的有效方法。如图 2 左侧所示, 给定一个大小为 9×9 的 SAIs 的输入失真 LFI, 首先提取其中心 $A \times A$ SAIs, 以降低计算复杂度。由于人眼对 YCbCr 颜色空间中的亮度分量 Y 更敏感, 因此仅将 Y 分量用于图像块生成。然后, 利用 $S \times S$ 滑动窗口以 T 步长在每个 SAI 上均匀滑动, 并将中心 $A \times A$ SAIs 的相同采样位置聚合并转换形成空间-角度图像块 I_{dis} 。具体可表示为

$$I_{CY} = \{rgb2ycbcr(Central(I))\}, \quad (1)$$

$$I_{dis} = Permute(Stich(ICY)), \quad (2)$$

其中: $I \in \mathbb{R}^{u \times v \times x \times y}$ 代表输入失真 LFI, $Central(\cdot)$ 代表取中心 $A \times A$ 的 SAIs, $rgb2ycbcr(\cdot)$ 代表只取 Y 通道分量, $Permute(Stich(\cdot))$ 代表移位切块并重组。在实验中, A, S, T 分别设置为 7、32、32。

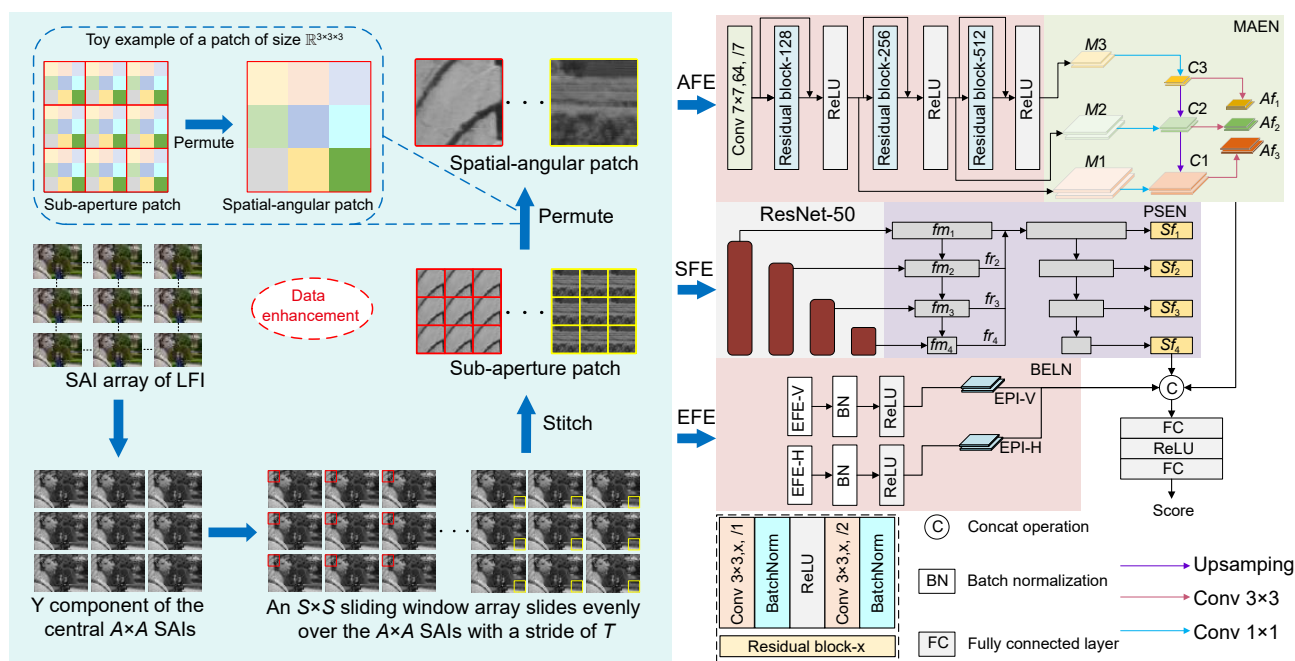


图2 SAE-BLFI 整体框架

Fig. 2 Overall framework of SAE-BLFI

2.3 空间-角度分离模块

考虑到光场图像是同时包含场景中光线强度和角度信息的高维信号, 本文采用一种空间-角度分离操作, 将原始的高维图像特征转换为低维图像特征, 以微透镜图像形式进行特征提取。具体来说, 通过对微透镜图像块使用两种不同的卷积操作, 实现角度域和空间域的分离。如图3所示, 为方便可视化, 示例为角度分辨率为 3×3 的微透镜图像块。其中来自不同视图的像素用数字表示, 来自不同宏像素的像素用不同颜色表示。在实验实施当中, 角度分辨率为 7×7 。

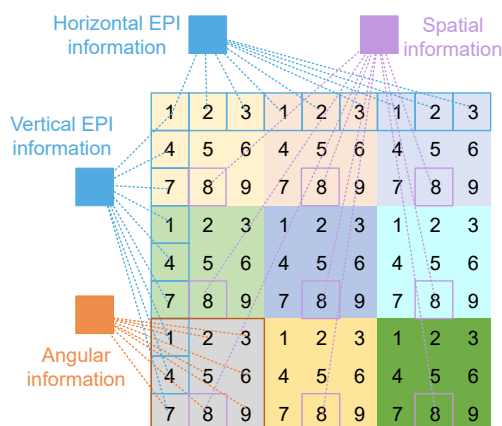


图3 空间-角度分离示意图

Fig. 3 Schematic diagram of spatial-angular separation

空间信息获取: 为了实现从图像块中获取到空间信息, 需要针对来自相同角度坐标的像素进行卷积,

来自其它角度视图的像素应该被隔离。为了实现该目标, 这里使用了卷积核大小为 3×3 、步幅大小为1、膨胀率为7的卷积。其中, 膨胀率设置为7是为了保证隔离来自不同视图的信息。

角度信息获取: 角度信息的提取需要保证来自同一空间坐标下的像素被卷积, 而隔离来自不同宏像素的其它像素。因此, 这里采用了卷积核大小为 7×7 、步幅大小为7、膨胀率为1的卷积。卷积核的大小以及步幅的选择是根据图像块的属性确定。上述图像块包含 7×7 的角度像素信息, 即包含来自同一个宏像素的所有像素信息。

2.4 空间-角度特征提取网络

在图像质量评价研究中, 增强特征的表示能力是至关重要的。多尺度特征由于能够通过不同尺度上的信息来更为全面地反映信号特征, 广泛应用于计算机视觉和深度学习领域^[20]。由于LFI不同于传统图像, 除了常见的空间失真以外, 还包含有复杂的角度失真。因此, 本文基于多尺度特征提取方式提出两种捕捉空间和角度失真的网络结构。

2.4.1 渐进式空间特征提取网络

基于对人类视觉系统 (human visual system, HVS) 的考虑, 本小节围绕卷积的特性设计一种渐进式空间特征提取网络 (progressive spatial feature extraction network, PSEN), 模拟由粗到细的学习方式。整个分

支主要由两个要素组成: 特征提取和深度迭代。

特征提取阶段利用 ResNet50 作为骨干网络框架, 并使用 ImageNet 中的预训练模型初始化相应的网络参数, 再分别获取来自不同层级的四个特征 fm_1, fm_2, fm_3, fm_4 ; HVS 的神经反馈作用被证明可以对失真区域增强细节表达, 因此, 在深度迭代阶段, 以下一深度阶段的特征作为反馈输入传递到当前阶段。通过三次迭代, 整合并关联来自不同深度的特征信息。迭代过程表示如下:

$$f_{td} = \gamma_d[(f_{md}, f_{md} - 1), \vartheta_d], \quad d = 2, 3, 4, \quad (3)$$

其中: f_{td} 表示深度为 d 时的输出, $f_{md} - 1$ 表示上一个深度的输出, ϑ_d 表示变换参数, γ_d 表示映射函数。

随后, 将所有的迭代输出与第一层级的特征进行融合, 送入到最后的卷积阶段得到四个空间特征块 $Sf_i (i = 1, 2, 3, 4)$, 经过卷积操作和平均池化之后获取到图像的空间信息 SF 。上述操作可表示为

$$SF = \text{Concat}\{Sf_1, Sf_2, Sf_3, Sf_4\}. \quad (4)$$

2.4.2 多尺度金字塔角度特征提取网络

角度信息是 LFI 区别于传统 2D 图像最显著的特点。因此, 有效捕捉到 LFI 中存在的细微角度失真决定 LF-IQA 性能好坏的关键要素。基于角度特性的考虑, 为实现角度域中的特征多重利用能力, 本节构建了一种多尺度角度金字塔特征提取网络 (multiscale angle pyramid feature extraction network, MAEN)。通过获取不同层次的失真信息, 增强特征的表达能力。其后, 通过一种类似金字塔的结构设计, 重复使用低层级和高层级特征信息, 模拟人类视觉的分级感知机制, 进一步加强多尺度特征的表达能力。该结构的具体细节如下: 首先设计 3 个残差块, 每个残差块包含两个卷积层、两个归一化层和一个 ReLU 激活函数。在每个卷积层之后使用 ReLU 激活函数以避免不良初始化。随后通过选择不同尺度的特征图实现多尺度特征融合, 其中选择了将三个残差块的输出作为融合选项。为了保证通道维度一致, 通过上采样操作对其分辨率进行加倍。最后使用卷积操作获取最终的特征图。该流程表示如下:

$$Af_1 = \xi(\psi(M1) \oplus \delta(\xi(\psi(M2) \oplus \delta(\psi(M3))))), \quad (5)$$

$$Af_2 = \xi(\psi(M2) \oplus \delta(\psi(M3))), \quad (6)$$

$$Af_3 = \xi(\psi(M3)), \quad (7)$$

其中: $\psi(\cdot)$ 表示 1×1 卷积操作, $\xi(\cdot)$ 表示 3×3 卷积操作, $\delta(\cdot)$ 表示上采样操作, $\oplus$ 表示相加操作。

2.5 双向 EPI 特征学习网络

EPI 可以通过记录不同视点位置下的光线强度分布来反映光场图像的角度特性。如图 4 所示, 左侧展示了来自 Win5-LID 数据集的两个示例 LFI (bikes、dishes), 右侧截取了来自原始参考图像以及四种失真图像的 EPI。其中, 四种失真分别是 JPEG、HEVC 两种压缩失真和线性插值 (LiNear interpolation, LN)、最近邻插值 (nearest neighbor interpolation, NN) 两种重建失真。从图中可以观察到同一场景中的不同失真所表现出来的 EPI 存在明显差异。

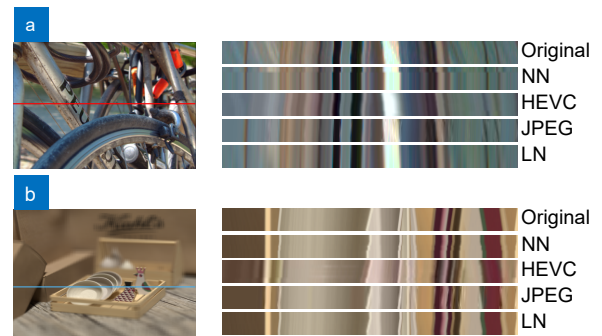


图 4 两个场景不同失真情况下的 EPI

Fig. 4 EPI under different distortion conditions for two scenarios

传统的方式一般利用 EPI 的梯度变化来捕捉角度失真。本文则是利用基于 CNN 的双向 EPI 特征学习网络 (bidirectional EPI feature learning network, BELN) 来直接提取 EPI 图像特征, 以获取来自水平角度和垂直角度上的失真细节信息。根据卷积的性质, 设计不同类型的卷积核, 形成不同的特征感受野, 捕捉水平和垂直方向的 EPI 图像信息。对于水平方向 EPI 特征获取, 需要固定 u 和 x 坐标, 以此获取来自 v - y 空间内的图像信息。因此使用一种特殊卷积方式: 卷积核大小为 1×7^2 , 水平步长为 7, 垂直步长为 1, 不进行填充操作。逐行卷积, 即可实现在每一个水平方向获取像素信息。垂直方向 EPI 获取采取类似方式: 卷积核大小为 $7^2 \times 1$, 水平步长为 1, 垂直步长为 7, 不进行填充操作。逐列卷积, 依次获取每一个垂直方向上的像素信息。

2.6 质量回归

将所提取到的空间特征、角度特征以及极平面特征进行级联, 经过线性回归输出得到每一个空间-角度块的预测质量分数 Q_i 。最终, 合并所有空间-角度块的质量分数作为整幅失真 LFI 的整体质量分数, 表

示为

$$QS = \frac{1}{N} \sum_{i=1}^N Q_i, i = 1 \cdots N, \quad (8)$$

其中: N 为单张 LFI 切块数量, QS 表示整幅失真 LFI 的感知质量分数。

3 实验结果与分析

为验证所提出 SAE-BLFI 的性能, 将其与具有代表性的 2D-IQA 和 LF-IQA 方法在三个公开光场图像数据集 Win5-LID^[21], NBU-LFI.0^[22], SHU^[23] 上进行实验。对比方法包括两种 2D NR-IQA 方法 (BRISQUE^[24], GWH-GLBP^[25])、3 种 FR LF-IQA 方法 (MD - FM^[5], Min 方法^[7], Meng 方法^[26]), 以及 7 种 NR LF-IQA 方法 (BELIF^[27], NR-LFQA^[8], Tensor-NLFQ^[12], VBLFI^[10], 4D-DCT-LFIQA^[2], DeeBLiF^[18] 和 SATV-BLiF^[28])。

3.1 光场图像数据集

Win 5-LID 数据集包含 6 个真实场景和 4 个合成场景作为参考 LFI。然后, 将六种失真类型应用于 10 幅参考 LFI, 共计产生 220 幅失真 LFI。这六种失真类型包括 HEVC 和 JPEG 2000 这两种压缩失真, 以及线性插值 (LN)、最近邻插值 (NN)、基于 CNN 的插值 (EPICNN 和 USCD) 这四种角度插值重建失真。除了两个 CNN 失真之外, 其它四个失真中的每一个都包含 5 个失真级别。主观实验采用双刺激连续质量量表法 (double-stimulus continuous quality scale, DSCQS) 在 5 点离散量表上评估这 220 幅 LFI 的质量。对于每幅 LFI, 收集平均意见评分 (mean opinion score, MOS)。

NBU-LFI.0 数据集包含 14 幅参考图像 (包括真实场景和合成场景) 和 210 幅失真的 LFI, 涵盖 5 种类型的重建失真, 包括最近邻插值 (NN)、双三次插值 (bicubic interpolation, BI)、基于学习的重建 (EPICNN)、基于视差图的重建插值 (MDR) 和空间超分辨率重建 (VDSR)。每个失真有 3 个级别。采用被动式结合交互式的可视化方法、结合 DSCQS 的方法进行主观实验。提供了 5 点离散标度的 MOS。

SHU 数据集包括来自 EPFL 数据集^[29] 的 8 幅源 LFI。每幅 LFI 都经过五种失真处理, 即高斯模糊、JPEG2000、JPEG、运动模糊和白噪声伪影, 共计生成了不同失真程度的 240 幅失真图像。主观实验采用被动实验模式和 DSCQS 方法。数据集提供的 MOS

范围为 0 (差) 至 5 (优)。

3.2 实施细节以及性能评价标准

所提出的方法在 PyTorch 框架上实现, 实验环境设置为 (CPU: Intel (R) Core (TM) i9-10900X CPU 3.70 GHZ; GPU: NVIDIA GeForce RTX 3090 24 G)。模型参数使用 mini-batch 随机梯度下降 (SGD) 优化器进行优化。权重动量和衰减率分别设置为 0.9 和 0.001。初始学习率为 0.001, 每 10 个 epochs 乘以 0.1, 共训练 20 个 epoch。使用批量大小为 32 进行从头训练。

现有的大多数用于质量评估的 LFI 数据集只包含由少量参考 LFI 生成的几百个失真 LFI。如果将整个数据集随机分成训练集和测试集, 这两个集合中可能会包含来自相同参考 LFI 的失真 LFI。因此, 为了确保训练集和测试集完全独立, 本文使用 K 折交叉验证^[18]来进行实验。具体来说, 对于每个数据集, 所有从同一参考场景生成的失真 LFI 被视为一个子集。这样, 一个包含 K 个参考场景的 LFI 数据集将被分成 K 个失真 LFI 子集, 每个子集仅包含一个参考场景的所有失真版本。然后, 随机选择 $K-2$ 个子集用于训练, 并在剩余的 2 个子集上测试评价方法的性能表现。共有 $K(K-1)/2$ 种可能的组合。最终, 基于所有可能的组合进行实验, 并将平均结果作为模型性能表现。

使用三个指标来衡量各种图像质量评价方法的性能: Pearsons 线性相关系数 (Pearson linear correlation coefficient, PLCC), Spearman 秩序相关系数 (Spearman rank order correlation coefficient, SROCC) 和均方根误差 (root mean square error, RMSE)。PLCC 测量客观评价方法预测结果与 MOS 值的线性关系, SROCC 关注单调性, RMSE 评估预测精度。较高的 PLCC 和 SROCC 值表示更好的性能, 而 RMSE 则相反。根据文献^[30]建议, 计算 PLCC 和 RMSE 之前采用了一个包含 5 个参数的非线性函数进行得分映射, 如式 (9) 所示:

$$f(p) = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + e^{\beta_2(p - \beta_3)}} \right) + \beta_4 p + \beta_5, \quad (9)$$

其中: $\beta_i (i = 1, \dots, 5)$ 是拟合参数, p 表示目标预测, $f(p)$ 是其非线性映射结果。

3.3 总体性能比较

在实验中, 基于学习的方法都使用 K 折交叉验证进行训练。为保证性能对比的公平性, 基于非学习

的方法也在相同的测试集上实验。对于所有对比方法, 均使用其作者发布的代码和默认参数设置, 以避免偏差。

表 1 展示了所提出方法和对比方法在 Win5-LID, NBU-LF1.0 以及 SHU 数据集上的整体性能对比, 最佳性能用粗体表示。从表 1 中可以发现, 传统的 2D IQA 方法都无法精准地预测 LFI 的质量。这主要是因为这些方法都没有考虑到 LFI 的角度失真对于人眼视觉感知的影响。而 FR/NR LF-IQA 方法基于 LFI 的特性出发, 额外考虑了角度失真对于光场图像质量的影响, 从而实现了更高的预测准确性。其中, 基于深度学习框架的 DeeBLiF 方法的性能明显优于其它基于传统手工特征的方法, 这可能是因为 CNN 对于图像特征的强大获取能力。然而, 大多数现有的 LF-IQA 方法在获取角度特征和空间特征时, 忽略了中间过程的特征, 只保留了最后尺度的信息, 所获取到的细节信息不全。而本文通过逐级获取图像特征, 充分利用多尺度特征, 更全面地反映失真信息, 并结合 EPI 进行特征补充, 增强角度特征的表达, 达到了相对更好的性能。在三个数据集上都取得最佳性能, 其中在 Win5-LID 和 NBU-LF1.0 数据集中 SROCC 表现出明显的优势。其原因可能是由于 Win5-LID 和 NBU-LF1.0 数据集包含更多的重建失真, 而 SHU 数据集主要包括 2D 图像中的常见失真。

由于实验采用的是 K 折交叉验证, 不同的划分

方式会导致训练集和测试集不同, 进而产生不同的结果。为了验证所提出的方法和其它 LF-IQA 方法在整体性能方面的稳定性, 图 5 展示了 Win5-LID 和 NBU-LF1.0 数据集上 SROCC 值的箱型图。在图 5 中, SROCC 值的中位值由红线表示。对于一种 IQA 方法而言, 其箱型图中心红线越高、箱体高度越短, 其预测性能和稳定性越好。

可以发现, 本文所提出的方法在性能稳定性上面要优于其它对比的 LF-IQA 方法。

3.4 针对不同失真类型的鲁棒性

为了进一步分析所提出的模型与其它 IQA 模型在不同失真类型下的性能, 在 Win5-LID 和 NBU-LF1.0 数据集上进行了单独失真类型的性能比较实验。这里不考虑 Win5-LID 数据集中基于 CNN 的两个失真类型, 因为它们只存在一个失真级别。表 2 中列出了 SROCC 的结果, 其中最佳性能和次佳性能分别用粗体和下划线显示。Hit 表示每种模型在各失真类型中检测性能最佳或次佳的次数。

表 2 显示, 所提出的 SAE-BLiF 在多数失真类型中取得了最佳性能或者次佳性能, 而且在个别失真类型 (如 JPEG2000) 上有显著的改善。对于大多数 FR/RR LF-IQA 方法, 由于参考了原始 LFI 的图像信息, 所以会呈现更好的鲁棒性。在没有参考图像信息参与的情况下, 2D-IQA 方法对压缩失真的性能明显优于对角度重建失真的性能, 而 LF-IQA 方法则是没

表 1 不同 LFI 数据集上不同方法的总体性能比较

Table 1 Overall performance comparison of different methods on different LFI datasets

Types	Methods	Win5-LID			NBU-LF1.0			SHU		
		PLCC↑	SROCC↑	RMSE↓	PLCC↑	SROCC↑	RMSE↓	PLCC↑	SROCC↑	RMSE↓
NR 2D-IQA	BRISQUE ^[24]	0.6217	0.4537	0.7604	0.4989	0.3871	0.7879	0.9011	0.8883	0.4591
	GLBP ^[26]	0.5357	0.4150	0.8130	0.5056	0.3490	0.7647	0.7168	0.6565	0.7504
FR LF-IQA	MDFM ^[6]	0.7763	0.7471	0.6249	0.7888	0.7559	0.5649	0.8947	0.8908	0.4863
	Min's ^[7]	0.7281	0.6645	0.6874	0.7104	0.6579	0.6439	0.8497	0.8470	0.5757
	Meng's ^[26]	0.6983	0.6347	0.7203	0.8404	0.7825	0.4889	0.9279	0.9203	0.4039
	BELIF ^[27]	0.5751	0.5059	0.7865	0.7014	0.6389	0.6276	0.8967	0.8656	0.4803
NR LF-IQA	NR-LFQA ^[8]	0.7298	0.6979	0.6271	0.8528	0.8113	0.4658	0.9224	0.9229	0.4132
	Tensor-NLFQ ^[12]	0.5813	0.4885	0.7706	0.6884	0.6246	0.6305	0.9307	0.9061	0.3857
	VBLFI ^[10]	0.7213	0.6704	0.6843	0.8027	0.7539	0.5218	0.9235	0.8996	0.4064
	4D-DCT-LFIQA ^[2]	0.8234	0.8074	0.5446	0.8395	0.8217	0.4871	0.9400	0.9320	0.3691
	DeeBLiF ^[18]	0.8427	0.8186	0.5160	0.8583	0.8229	0.4588	0.9548	0.9419	0.3185
	SATV-BLiF ^[28]	0.7933	0.7704	0.5842	0.8515	0.8237	0.4686	0.9332	0.9284	0.3897
	Proposed	0.8653	0.8451	0.4863	0.9108	0.8937	0.3658	0.9649	0.9547	0.2808

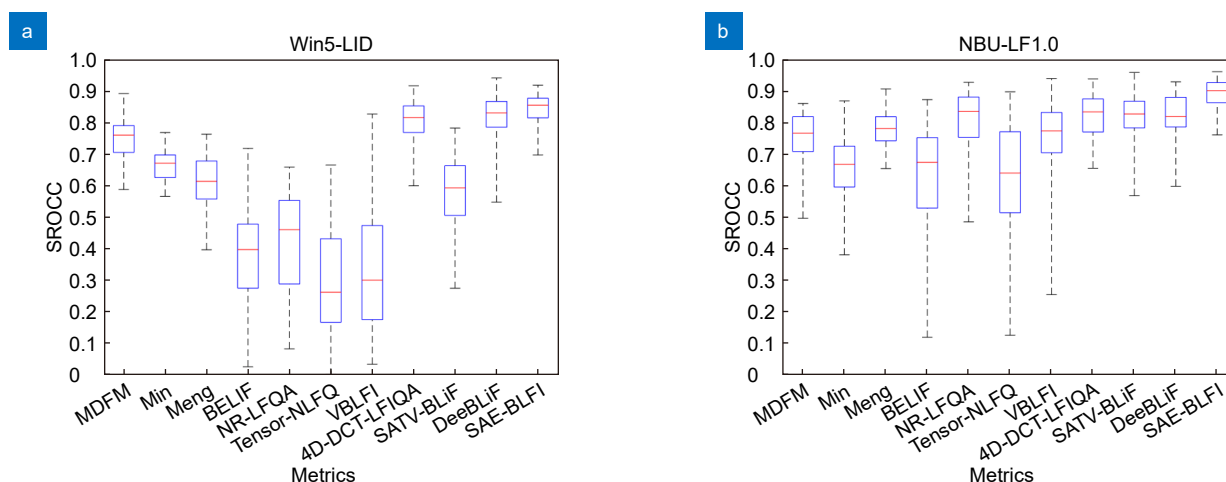


图 5 Win5-LID 和 NBU-LF1.0 数据集 K 折交叉验证中 SROCC 分布箱型图。(a) Win5-LID; (b) NBU-LF1.0

Fig. 5 Boxplot of SROCC distribution in K-fold cross-validation on Win5-LID and NBU-LF1.0 datasets. (a) Win5-LID; (b) NBU-LF1.0

表 2 在 Win5-LID 和 NBU-LF1.0 数据集上, 不同方法针对于不同失真类型的 SROCC 值

Table 2 SROCC values for different distortion types across various methods on Win5-LID and NBU-LF1.0 datasets

Types	Methods	Win5-LID				NBU-LF1.0					Hit count
		HEVC	JEPG2K	LN	NN	NN	BI	EPICNN	MDR	VDSR	
NR 2DIQA	BRISQUE ^[24]	0.5641	0.7801	0.5222	0.2462	0.3435	0.4145	0.5795	0.4331	0.7937	0
	GLBP ^[25]	0.7165	0.4853	0.4678	0.3011	0.3229	0.3995	0.4344	0.4478	0.7381	0
	MDFM ^[5]	0.7922	0.7669	0.6437	0.6692	0.8025	0.9089	0.7899	0.7386	0.8709	1
FR LF-IQA	Min's ^[7]	0.6997	0.6507	0.6159	0.6288	0.8156	0.8667	0.7361	0.7963	0.9376	1
	Meng's ^[26]	0.8886	0.6939	0.8459	0.8001	0.7429	0.9018	0.7997	0.5783	0.9225	2
	BELIF ^[27]	0.7666	0.6379	0.6097	0.5452	0.7680	0.7122	0.6874	0.6128	0.7989	0
NR LF-IQA	NR-LFQA ^[8]	0.7571	0.7338	0.6362	0.7026	0.8930	0.8807	0.7653	0.6111	0.8164	0
	Tensor-NLFQ ^[12]	0.6853	0.5799	0.5663	0.5897	0.6946	0.7203	0.5245	0.5417	0.8018	0
	VBLFI ^[10]	0.7141	0.7449	0.6908	0.7197	0.8316	0.8372	0.7195	0.4613	0.9134	0
	4D-DCT-LFQA ^[2]	0.8698	0.8946	0.8127	0.8235	0.9040	0.8719	0.7100	0.8095	0.8882	2
	DeeBLiF ^[18]	0.9648	0.8195	0.7928	0.8306	0.9184	0.8876	0.7248	0.6961	0.8857	3
	SATV-BLiF ^[28]	0.7918	0.8685	0.7566	0.8525	0.9282	0.9190	0.7722	0.6498	0.8617	2
	Proposed	0.9417	0.8955	0.8472	0.8742	0.9165	0.9153	0.7749	0.8443	0.9294	7

有明显变化, 主要是因为 2D-IQA 方法并没有考虑 LFI 的角度一致性。这表明角度一致性特征对于预测整体 LFI 质量有很大帮助。最终实验结果表明, 本文提出的方法在面对不同的失真类型时具有较好的鲁棒性。

3.5 消融实验

针对所提方法中的三种特征提取网络进行了消融实验, 包括渐进式空间特征提取网络、多尺度金字塔角度特征提取网络和双向 EPI 特征学习网络。表 3 展示了在 Win5-LID 和 NBU-LF1.0 数据集上的消融实验的具体结果, 其中, 最佳性能用粗体表示, SF 代表

空间特征, AF 代表角度特征, EF 代表 EPI 特征。当只进行简单的空间或者角度特征提取都会导致获取到的失真信息不完整, 从而影响对失真图像质量的准确预测。当获取图像的空间特征信息和角度特征信息的同时增强对角度一致性考虑时, 性能达到最佳。

3.6 时间复杂度分析

对本文方法与七种无参考光场图像质量评价方法进行了时间复杂度对比实验。为了公平比较, 所有指标在相同的硬件配置 (CPU: Intel (R) Core (TM) i9-10900X CPU 3.70 GHz; GPU: NVIDIA GeForce RTX 3090 24 G; 内存: 64 G RAM) 下执行。测试时

间为评价单幅光场图像的运行时间, 不包括数据加载和模型初始化的时间。所有基于 CNN 的方法均使用 Pytorch 进行测试, 而其它方法使用 Matlab 进行测试。表 4 展示了这些方法在 Win5-LID 数据集中 Flowers_HEVC_44 场景上测试时间的对比结果, 从中可以看出, 大多数传统 NR LF-IQA 方法耗时较长, 因为它们依赖于提取多个手工特征以确保预测准确性。基于深度学习的方法中, 本文提出的方法虽然略慢于 DeeBLiF^[18], 但如表 1 所示取得了更加精准的预测效果。

3.7 跨库实验

为证明所提方法在不同数据集上同样存在良好的泛化性, 本小节在不同数据集进行跨库验证。Win5-LID 和 NBU-LF1.0 数据集中都包含 NN 失真, 并且 Win5-LID 和 SHU 数据集中都包含 JPEG2000 失真。

因此, 在进行跨数据集验证时选择了在 Win5-LID 数据集进行训练, 在 NBU-LF1.0 数据集和 SHU 数据集上进行测试。测试结果如表 5 所示, 可以发现本文所提出的方法在不同的数据集上进行测试时仍然具有良好的性能。

3.8 统计显著性检测

为了测试任意两种 LF-IQA 方法之间的显著性关系, 采用了 F-test 对任意两种方法进行测试。具体操作是利用 Matlab 中的 `vartest2` 函数进行计算, 置信水平设置为 95%。图 6 是分别在 Winds-LID 和 NBU-LF1.0 数据集测试的结果。其中“1”、“0”和“-1”分别代表行方法要比列方法具有更好的统计性能、二者不易区分、行方法比列方法具有更差的统计性能。从图中可以看出, 本文提出的方法明显优于其它 LF-IQA 方法。

表 3 Win5-LID 和 NBU-LF1.0 数据集上不同功能模块的消融实验

Table 3 Ablation experiments of different functional modules on Win5-LID and NBU-LF1.0 datasets

	Win5-LID			NBU-LF1.0		
	PLCC	SROCC	RMSE	PLCC	SROCC	RMSE
SF	0.8338	0.8179	0.5217	0.8853	0.8654	0.4104
AF	0.8195	0.7992	0.5322	0.8709	0.8566	0.4208
EF	0.7950	0.7992	0.5652	0.8637	0.8478	0.4176
SF+AF	0.8513	0.8285	0.5057	0.9057	0.8857	0.3845
SF+AF+EF	0.8653	0.8451	0.4863	0.9108	0.8937	0.3658

表 4 不同 NR LF-IQA 方法运行时间对比

Table 4 Comparison of running time for different NR LF-IQA methods

Methods	Platform	Device	Time/s
BELiF ^[27]	Matlab	CPU	167.60
NR-LFQA ^[8]	Matlab	CPU	220.92
Tensor-NLFQ ^[12]	Matlab	CPU	630.47
VBLFi ^[10]	Matlab	CPU	68.48
4D-DCT-LFIQA ^[2]	Matlab	CPU	148.29
DeeBLiF ^[18]	Pytorch	GPU	2.77
SATV-BLiF ^[28]	Matlab	CPU	4.38
Proposed	Pytorch	GPU	3.77

表 5 在 Win5-LID 数据集上训练模型并在 NBU-LF1.0 和 SHU 数据集上测试的结果

Table 5 The results of training the model on the Win5-LID dataset and testing it on the NBU-LF1.0 and SHU datasets

Methods	NBU-LF1.0 (NN)		SHU (JPEG2000)	
	PLCC	SROCC	PLCC	SROCC
4D-DCT-LFIQA	0.7753	0.7040	0.7824	0.7967
DeeBLiF	0.8253	0.7265	0.7609	0.7252
Proposed	0.9082	0.8610	0.8821	0.8717

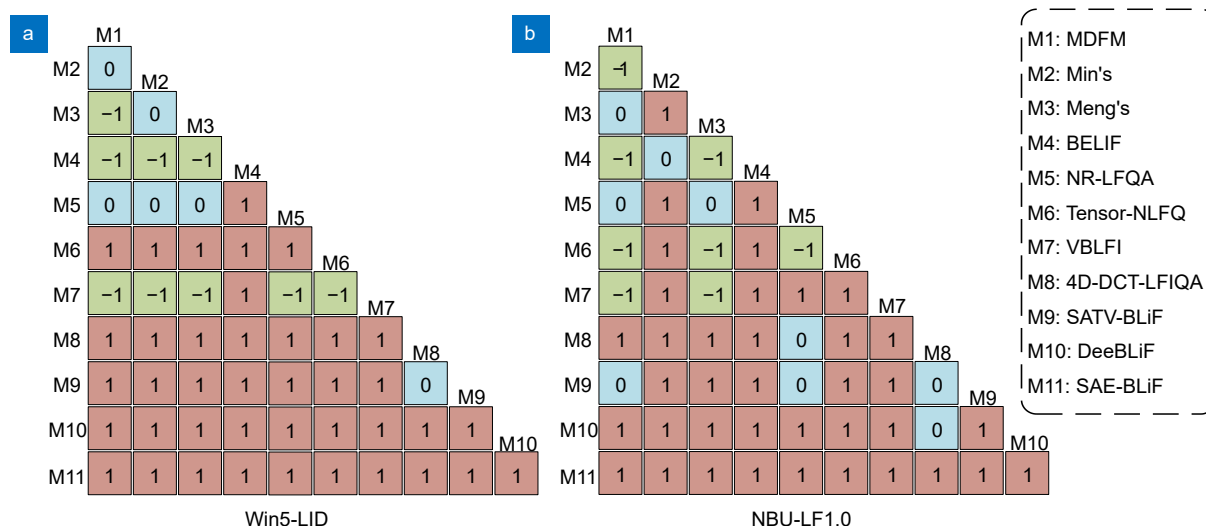


图6 Win5-LID 和 NBU-LF1.0 数据集的 F 检验统计显著性分析。(a) Win5-LID; (b) NBU-LF1.0

Fig. 6 F-test statistical significance analysis on Win5-LID and NBU-LF1.0 datasets. (a) Win5-LID; (b) NBU-LF1.0

4 结论

本文提出一种有效的空角感知测量的无参考光场图像评价框架, 利用图像块的预测质量来预估整幅失真图像的图像质量。针对 LF-IQA 任务中数据量不足的问题, 通过数据增强操作对数据进行扩充, 并同时实现空间-角度图像块的生成。为了模拟人眼感知特性, 设计两种多尺度特征提取方式构建局部特征与全局特征的细微关联。通过渐进式空间特征提取网络获取失真 LFI 的空间失真信息; 而在角度失真衡量时, 则是采用多尺度金字塔的结构实现, 从而完成失真 LFI 空间失真和角度失真的信息获取。考虑到 LFI 独特的角度特性, 额外设计了双向 EPI 特征学习网络来增强对于光场图像角度一致性的考量。在三个公开光场图像质量评价库上的对比实验结果表明, 本文提出的方法明显优于当前经典的 2D/LF-IQA 方法, 达到了更精准的预测效果。当前的工作大多是围绕图像的空域信息展开, 未来的工作中将探索频域与空域结合对 LF-IQA 带来的影响。

利益冲突: 所有作者声明无利益冲突

参考文献

- [1] Zuo C, Chen Q. Computational optical imaging: an overview[J]. *Infrared Laser Eng*, 2022, **51**(2): 20220110.
左超, 陈钱. 计算光学成像: 何来, 何处, 何去, 何从?[J]. *红外与激光工程*, 2022, **51**(2): 20220110.
- [2] Xiang J J, Jiang G Y, Yu M, et al. No-reference light field image quality assessment using four-dimensional sparse transform[J]. *IEEE Trans Multimedia*, 2023, **25**: 457–472.
- [3] 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.
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.
- [4] Yu M, Liu C. Single exposure light field imaging based all-in-focus image reconstruction technology[J]. *J Appl Opt*, 2021, **42**(1): 71–78.
于森, 刘诚. 基于单次曝光光场成像的全焦图像重建技术[J]. *应用光学*, 2021, **42**(1): 71–78.
- [5] Tian Y, Zeng H Q, Xing L, et al. A multi-order derivative feature-based quality assessment model for light field image[J]. *J Vis Commun Image Represent*, 2018, **57**: 212–217.
- [6] Huang H L, Zeng H Q, Hou J H, et al. A spatial and geometry feature-based quality assessment model for the light field images[J]. *IEEE Trans Image Process*, 2022, **31**: 3765–3779.
- [7] Min X K, Zhou J T, Zhai G T, et al. A metric for light field reconstruction, compression, and display quality evaluation[J]. *IEEE Trans Image Process*, 2020, **29**: 3790–3804.
- [8] Shi L K, Zhou W, Chen Z B, et al. No-reference light field image quality assessment based on spatial-angular measurement[J]. *IEEE Trans Circuits Syst Video Technol*, 2020, **30**(11): 4114–4128.
- [9] Luo Z Y, Zhou W, Shi L K, et al. No-reference light field image quality assessment based on micro-lens image[C]//2019 Picture Coding Symposium (PCS), 2019: 1–5.
<https://doi.org/10.1109/PCS48520.2019.8954551>.
- [10] Xiang J J, Yu M, Chen H, et al. VBLFI: visualization-based blind light field image quality assessment[C]//2020 IEEE International Conference on Multimedia and Expo (ICME), 2020: 1–6. <https://doi.org/10.1109/ICME46284.2020.9102963>.
- [11] Lamichhane K, Battisti F, Paudyal P, et al. Exploiting saliency in quality assessment for light field images[C]//2021 Picture Coding Symposium (PCS), 2021: 1–5.
<https://doi.org/10.1109/PCS50896.2021.9477451>.
- [12] Zhou W, Shi L K, Chen Z B, et al. Tensor oriented no-reference light field image quality assessment[J]. *IEEE Trans Image Process*, 2020, **29**: 4070–4084.

- [13] Xiang J J, Yu M, Jiang G Y, et al. Pseudo video and refocused images-based blind light field image quality assessment[J]. *IEEE Trans Circuits Syst Video Technol*, 2021, 31(7): 2575–2590.
- [14] Alamgeer S, Farias M C Q. No-Reference light field image quality assessment method based on a long-short term memory neural network[C]//2022 *IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, 2022: 1–6. <https://doi.org/10.1109/ICMEW56448.2022.9859419>.
- [15] Qu Q, Chen X M, Chung V, et al. Light field image quality assessment with auxiliary learning based on depthwise and anglewise separable convolutions[J]. *IEEE Trans Broadcast*, 2021, 67(4): 837–850.
- [16] Qu Q, Chen X M, Chung Y Y, et al. LFACon: Introducing anglewise attention to no-reference quality assessment in light field space[J]. *IEEE Trans Vis Comput Graph*, 2023, 29(5): 2239–2248.
- [17] Zhao P, Chen X M, Chung V, et al. DeLFIQE: a low-complexity deep learning-based light field image quality evaluator[J]. *IEEE Trans Instrum Meas*, 2021, 70: 5014811.
- [18] Zhang Z Y, Tian S S, Zou W B, et al. Deeblif: deep blind light field image quality assessment by extracting angular and spatial information[C]//2022 *IEEE International Conference on Image Processing (ICIP)*, 2022: 2266–2270. <https://doi.org/10.1109/ICIP46576.2022.9897951>.
- [19] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770–778. <https://doi.org/10.1109/CVPR.2016.90>.
- [20] Lu S Y, Ding Y M, Liu M Z, et al. Multiscale feature extraction and fusion of image and text in VQA[J]. *Int J Comput Intell Syst*, 2023, 16(1): 54.
- [21] Shi L K, Zhao S Y, Zhou W, et al. Perceptual evaluation of light field image[C]//2018 *25th IEEE International Conference on Image Processing (ICIP)*, 2018: 41–45. <https://doi.org/10.1109/ICIP.2018.8451077>.
- [22] Huang Z J, Yu M, Jiang G Y, et al. Reconstruction distortion oriented light field image dataset for visual communication[C]//2019 *International Symposium on Networks, Computers and Communications (ISNCC)*, 2019: 1–5. <https://doi.org/10.1109/ISNCC.2019.8909170>.
- [23] Shan L, An P, Meng C L, et al. A no-reference image quality assessment metric by multiple characteristics of light field images[J]. *IEEE Access*, 2019, 7: 127217–127229.
- [24] Mittal A, Moorthy A K, Bovik A C. No-reference image quality assessment in the spatial domain[J]. *IEEE Trans Image Process*, 2012, 21(12): 4695–4708.
- [25] Li Q H, Lin W S, Fang Y M. No-reference quality assessment for multiply-distorted images in gradient domain[J]. *IEEE Signal Process Lett*, 2016, 23(4): 541–545.
- [26] Meng C L, An P, Huang X P, et al. Full reference light field image quality evaluation based on angular-spatial characteristic[J]. *IEEE Signal Process Lett*, 2020, 27: 525–529.
- [27] Shi L K, Zhao S Y, Chen Z B. Belif: blind quality evaluator of light field image with tensor structure variation index[C]//2019 *IEEE International Conference on Image Processing (ICIP)*, 2019: 3781–3785. <https://doi.org/10.1109/ICIP.2019.8803559>.
- [28] Zhang Z Y, Tian S S, Zou W B, et al. Blind quality assessment of light field image based on spatio-angular textural variation[C]//2023 *IEEE International Conference on Image Processing (ICIP)*, 2023: 2385–2389. <https://doi.org/10.1109/ICIP49359.2023.10222216>.
- [29] Rerabek M, Ebrahimi T. New light field image dataset[C]//8th *International Conference on Quality of Multimedia Experience (QoMEX)*, 2016: 1–2.
- [30] VQEG. Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, 2000 [EB/OL]. <http://www.vqeg.org>.

作者简介



王斌(1999-), 男, 河南信阳人, 硕士研究生, 2022 年于湖北文理学院获得学士学位, 现为宁波大学信息科学与工程学院硕士研究生, 主要从事光场图像质量评价等方面的研究。

E-mail: wangbin4139@163.com



【通信作者】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(亚洲大学)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。

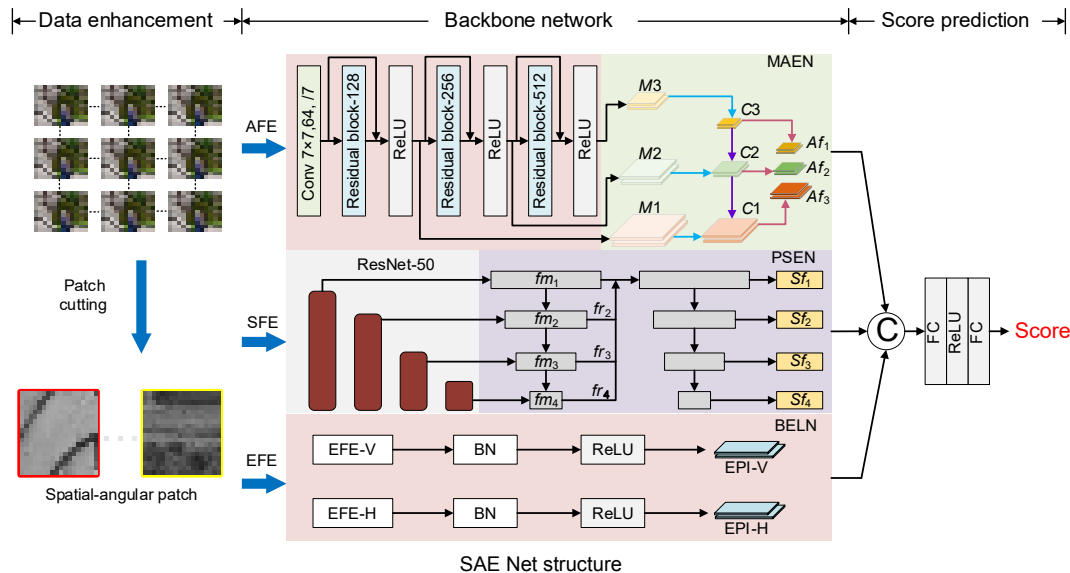
E-mail: yumei@nbu.edu.cn



扫描二维码, 获取PDF全文

No-reference light field image quality assessment based on joint spatial-angular information

Wang Bin¹, Bai Yongqiang², Zhu Zhongjie², Yu Mei^{1*}, Jiang Gangyi¹



Overview: Light field imaging, as an emerging media dissemination method, differs from traditional 2D and stereoscopic images in its ability to capture the intensity of light in scenes and the directional information of light rays in free space. Due to its rich spatial and angular information, light field imaging finds extensive applications in depth estimation, refocusing, and 3D reconstruction. However, during acquisition, compression, transmission, and reconstruction, light field images inevitably suffer from various distortions, leading to a decline in image quality. Light field image quality assessment (LFIQA) plays a crucial role in enhancing the quality of these images. Based on the characteristics of light field images, this paper proposes a no-reference image quality assessment (NRIQA) scheme that integrates spatial-angular information and epipolar plane image (EPI) information using deep learning. Specifically, this approach estimates the overall quality of distorted light field images by assessing the perceptual quality of image blocks. To simulate human visual perception, it employs two multi-scale feature extraction methods to establish subtle correlations between local and global features, thereby capturing information on spatial and angular distortions. Considering the unique angular properties of light field images, a bidirectional EPI feature learning network is additionally designed to acquire vertical and horizontal disparity information, enhancing consideration of angular consistency distortions in images. Finally, by aggregating across different features, the method integrates three distinct image features to predict the quality of distorted images. Experimental results conducted on three publicly available light field image quality assessment datasets demonstrate that the proposed method achieves higher consistency between objective quality prediction and subjective evaluation, showcasing excellent predictive accuracy.

Wang B, Bai Y Q, Zhu Z J, et al. No-reference light field image quality assessment based on joint spatial-angular information[J]. *Opto-Electron Eng*, 2024, 51(9): 240139; DOI: [10.12086/oe.2024.240139](https://doi.org/10.12086/oe.2024.240139)

Foundation item: Project supported by National Natural Science Foundation of China (62271276)

¹Faculty of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China; ²College of Information and Intelligent Engineering, Zhejiang Wanli University, Ningbo, Zhejiang 315100, China

* E-mail: yumei@nbu.edu