

## 基于多视觉特征聚合的光场质量评价方法

邹卓成, 邱钧, 刘畅\*

北京信息科技大学应用数学研究所, 北京 100101

**摘要** 从全光函数的空-角耦合关系出发, 分别从空间域、角度域、耦合域和投影域上提取视觉特征, 并提出一种无参考光场质量评价方法。在空间域上提取中心子孔径图的自然场景统计(NSS)特征; 在角度域和空-角耦合域上分别提取宏像素和极平面图(EPI)上的灰度共生矩阵(GLCM)特征; 在投影域上提取重聚焦图局部熵统计分布特征。最后将多视觉特征融合, 形成光场的视觉特征向量, 使用支持向量回归(SVR)训练出评分模型, 建立多视觉特征聚合的光场质量评价方法。实验结果表明, 所提评价方法对光场的评分与主观评分有较好的一致性。

**关键词** 图像处理; 光场; 质量评价; 特征提取; 支持向量回归

中图分类号 TP391

文献标志码 A

doi: 10.3788/AOS202141.1610002

## Light-Field Image Quality Assessment Based on Multiple Visual Feature Aggregation

Zou Zhuocheng, Qiu Jun, Liu Chang\*

*Institute of Applied Mathematics, Beijing Information Science & Technology University, Beijing 100101, China*

**Abstract** In this paper, we extract visual features from the spatial, angular, coupling, and projection domains considering the spatial-angular coupling of the plenoptic function to propose a nonreference method for assessing light-field image quality. Natural scene statistics (NSS) features of the central subaperture image are extracted in the spatial domain. In addition, macropixel and grayscale co-occurrence matrix (GLCM) features on epipolar plane images (EPI) are extracted in the angular and spatial-angular coupling domains, respectively, and local entropy statistical distribution characteristics of the refocusing maps are extracted in the projection domain. Then, multiple visual features are fused to form the visual feature vector of the light field, and support vector regression (SVR) is applied to train a scoring model. Thus, the light-field image quality assessment method based on multiple visual feature aggregation is established. Experimental results show that the proposed method shows consistency between the light-field score and subjective score.

**Key words** image processing; light field; quality evaluation; feature extraction; support vector regression

**OCIS codes** 100.2960; 110.1758; 110.3000

## 1 引言

视觉是人类的重要感知能力, 为了追求极致的视觉体验, 人类一直在探索视觉信息的采集与显示, 光场成像技术作为视觉信息探索中的一个重大成果, 引起了学术界和工业界的广泛关注。随之而来的是对光场数据处理的研究越来越多, 例如光场数

据的压缩<sup>[1]</sup>、重建<sup>[2]</sup>、渲染<sup>[3]</sup>处理等。这些处理操作都会引入不同程度的失真, 光场质量评价便是为了量化这些失真程度。准确的光场质量评价可以指导对光场数据处理方法的研究, 同时有助于分析光场成像系统的性能。

相比于传统视觉质量评价, 光场的视觉特征蕴含在高维光场数据中, 难以通过直接显示的方式进

收稿日期: 2021-01-28; 修回日期: 2021-03-08; 录用日期: 2021-03-18

基金项目: 国家自然科学基金(61931003, 61871042, 61801031)、北京市教委科研计划一般项目(KM201911232009)

通信作者: \*liuchang@bistu.edu.cn

行视觉评价。尽管目前有诸多关于光场显示技术的研究,例如传统光场显示<sup>[4]</sup>、多层光场显示<sup>[5]</sup>和压缩光场显示<sup>[6]</sup>,但是这些显示技术成本极高且效果欠佳,尚未得到广泛使用。现如今,对光场进行质量评价时主要是对光场降维后的视觉形式进行分析评价的。双平面参数化模型下的光场表示为四维函数  $L(u, v, s, t)$ <sup>[7]</sup>,固定其中任意 2 个坐标便可降维成 3 种不同形式的二维图像,分别为子孔径图像(SAI)、极平面图像(EPI)和微透镜图像(MLI)。除了二维视觉形式之外,还可通过计算光场数据获得重聚焦图像(RI)。光场数据的不同视觉形式可提供给人们不同的视觉信息,SAI 表示不同角度下对同一空间采样的结果,能反映出光场空间域上的清晰度变化;EPI 表示光场数据的空-角之间的关系,能反映光场空-角耦合域上的结构变化;MLI 表示场景空间中每个点沿不同角度方向的光线信息,能反映出光场角度域的一致性;RI 表示光场数据投影在不同成像平面的图像序列,能反映出光场在不同空间深度上的聚焦性能。光场质量评价主要通过分析这些降维后的视觉形式,定量表征光场的质量。

本文从提取光场数据所蕴含的视觉特征出发,提出一种无参考光场质量评价方法。评价指标主要体现在空间域清晰度、角度域一致性、空-角耦合域的耦合关系、投影域重聚焦特性上。通过有针对性地对光场数据的不同视觉形式进行视觉特征提取,得到相应特征量来表征光场在不同域上的质量变化情况。通过将所有视觉形式的特征量融合成全局特征向量,来表征光场整体质量,使用支持向量回归(SVR)<sup>[8]</sup>训练无参考回归模型映射全局向量所对应的质量分数。实验结果表明,所提方法训练出来的回归模型所预测的评分与主观评分有较高的一致性。

## 2 相关工作

近年来,国内外研究团队都开展了关于光场质量评价的相关研究。主观评价方面,通过将光场数据以不同的视觉形式展示给观察者,观察者进行主观打分<sup>[9-12]</sup>,以此来研究人类视觉系统(HVS)对光场数据的感知质量,并且建立了相应的光场质量评价数据集。主观评价虽然可靠,但耗时耗力,做不到实时评价,难以应用到各种实时处理系统中。

相比光场的主观质量评价,客观质量评价因可以嵌入到智能系统中实现实时评价而更具有实用性。近年来,Amirpour 等<sup>[13]</sup>为了评价光场数据压

缩后的质量,结合颜色空间信息,使用多种 2D 图像客观评价方法(PSNR-Y、PSNR-YUV 和 SSIM-Y 等)对光场子孔径图像进行了评价并且分析了评价的可靠性。Tian 等<sup>[14]</sup>提出多阶导数特征模型(MDFM),通过对比参考 SAI 和失真 SAI 之间的多阶导数差异来衡量光场质量失真程度。上述两种方案都只考虑了光场空间域变化情况,而没有考虑光场角度一致性退化情况。Fang 等<sup>[15]</sup>提出一种梯度幅值相似性方法,除了对比参考光场和失真光场在 SAI 上的梯度幅值相似性之外,还加入了 EPI 的对比来评价角度域上的一致性退化。Min 等<sup>[16]</sup>先通过对比失真 SAI 与参考 SAI 上的角点信息差异来衡量光场在整体空间上的质量,再通过计算边缘区域的均方误差(MSE)来衡量局部质量,最后通过分析多视图的质量来测量角度一致性退化。考虑到上述方法<sup>[13-16]</sup>所需参考信息量较大,Paudyal 等<sup>[17]</sup>仅从光场的深度信息出发,使用结构相似度(SSIM)<sup>[18]</sup>来度量参考光场与失真光场深度图的相似度,从而给出光场数据的质量。以上光场数据客观评价方法都是在有参考光场的前提下提出的,而实际应用中很多情况是无法提供参考光场信息的。在实际应用中光场质量的无参考评价对光场成像系统的研发、压缩算法的设计、计算成像算法的评估都具有重要的意义。

光场无参考质量评价一般是通过提取相关特征量来训练质量评价模型的。目前,Luo 等<sup>[19]</sup>通过提取 SAI 上的等价局部二进制(ULBP)特征和 MLI 上的熵特征来描述光场质量,再通过回归训练得出评价模型。Shan 等<sup>[20]</sup>则通过提取 SAI 各颜色通道上的均值、纹理特征及深度图上的特征来进行回归训练得出评价模型。这两种客观质量评价仅考虑了空间域和角度域上的信息,对于光场质量评价,还不够全面和准确。

为了更加准确地评价光场质量,本文基于光场数据所蕴含的视觉特征,分别从光场空间域的清晰度、角度域的一致性、空-角耦合域 EPI 上的空-角关系、投影域 RI 的重聚焦特性 4 个方面来综合衡量光场质量退化情况。对由微透镜阵列相机采集的光场数据进行实验,结果表明,所提方法得出的评价结果与主观评价具有较好的一致性。

## 3 光场多视觉特征量提取

对于光场数据  $L(u, v, s, t)$ ,分别从空间域( $s, t$ )、角度域( $u, v$ )、空-角耦合域( $u, s$ )( $v, t$ )、投影域

上提取视觉特征量来表征光场的质量,整体流程如图 1 所示。在空间域上,先计算中心子孔径图像的自然场景统计(NSS)<sup>[21]</sup>分布,然后通过使用非对称广义高斯分布(AGGD)<sup>[22]</sup>函数拟合出分布参数来描述空间清晰度的退化。在角度域上,先计算每个宏像素的灰度共生矩阵(GLCM)<sup>[23]</sup>,然后再提取 GLCM 的一致性和信息熵两个特征来描述角度一

致性退化。在空-角耦合域上,先计算 EPI 的 GLCM,然后提取 GLCM 的能量、对比度和一致性三个特征来描述空-角关系的结构性变化。在投影域上,通过提取重聚焦图的局部熵统计分布(SDLE)特征来描述重聚焦特性的变化。最后将所提取的各部分的特征量组合为一个向量,使用 SVR 进行回归训练,得出一种无参考综合评价模型。

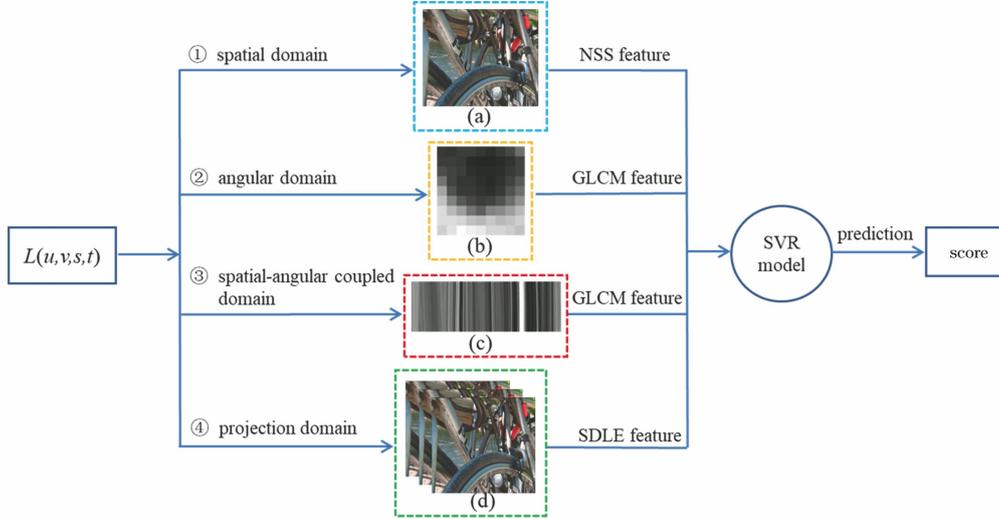


图 1 所提方法流程框图。(a)中心子孔径图;(b)宏像素图;(c)极平面图;(d)重聚焦图序列

Fig. 1 Framework of the proposed method. (a) Center view of the SAI; (b) macro-pixel image; (c) EPI; (d) RI sequence

### 3.1 空间域特征

光场的不同子孔径图像是不同视点下的二维透视投影图像,相当于从不同角度对场景空间进行采样的结果,表示为  $L_{u^*,v^*}(s,t)$ 。通常为了对光场的空间质量进行评价,需考虑所有子孔径图像的信息,但由于微透镜阵列型光场数据各视点图像之间的变化是像素级的,因此可忽略其他子孔径图像,仅使用中心子孔径图像上的特征便能反映出光场在空间域上的失真程度。

为了提取中心子孔径图中的 NSS 特征量<sup>[24]</sup>,先对中心子孔径图像进行灰度化,然后再使用经典的 NSS 模型对中心子孔径图像  $L_{u^*,v^*}(s,t)$  进行去均值对比度归一化(MSCN)处理,获取 MSCN 系数  $\hat{L}_{u^*,v^*}(s,t)$ ,具体计算公式为

$$\hat{L}_{u^*,v^*}(s,t) = \frac{L_{u^*,v^*}(s,t) - \mu_{u^*,v^*}(s,t)}{\sigma_{u^*,v^*}(s,t) + 1}, \quad (1)$$

式中: $u^*$ 和 $v^*$ 为中心子孔径图的坐标值; $s$ 和 $t$ 为图像的像素坐标; $\mu_{u^*,v^*}(s,t)$ 和 $\sigma_{u^*,v^*}(s,t)$ 分别为局部像素的均值和方差。

$$\mu_{u^*,v^*}(s,t) = \sum_{k=-K}^K \sum_{l=-L}^L \omega_{k,l} L_{u^*,v^*}^{k,l}(s,t), \quad (2)$$

$$\sigma_{u^*,v^*}(s,t) = \sqrt{\sum_{k=-K}^K \sum_{l=-L}^L \omega_{k,l} [L_{u^*,v^*}(s,t) - \mu_{u^*,v^*}(s,t)]^2}, \quad (3)$$

式中:二维高斯窗口  $\omega = \{\omega_{k,l} | k = -K, \dots, K; l = -L, \dots, L\}$ ,实验时选取  $K = L = 3$ 。

为了体现质量的变化情况,对  $\hat{L}_{u^*,v^*}(s,t)$  像素值进行统计。如图 2 所示,分别对比了不同场景光场在高效视频编码(HEVC)压缩<sup>[25]</sup>和 JPEG 压缩<sup>[1]</sup>两种失真情况下的中心子孔径图的 MSCN 系数统计分布,每种压缩失真方式包含 5 个失真等级,其后面数字越大,失真等级越高,则失真越严重<sup>[11]</sup>。第一行与第二行分别对比了同一场景不同失真方式下的 MSCN 系数分布,第一列与第二列分别对比了不同场景同一失真方式下的 MSCN 系数分布。从图 2 可以发现,分布曲线呈近似的正态分布,且随着光场失真变严重,统计分布曲线会变得更窄,说明中心子孔径图像能够较好地反映出光场空间域上的失真情况。为了提取 MSCN 系数分布的特征,使用 AGGD 函数对图 2 中的 MSCN 系数统计分布进行拟合,所求得

的参数作为特征量,具体过程的计算公式为

$$f(x; \alpha, \sigma_l^2, \sigma_r^2) = \begin{cases} \frac{\alpha}{(\beta_l + \beta_r)\Gamma(\frac{1}{\alpha})} \exp\left[-\left(\frac{-x}{\beta_l}\right)^\alpha\right], & x < 0 \\ \frac{\alpha}{(\beta_l + \beta_r)\Gamma(\frac{1}{\alpha})} \exp\left[-\left(\frac{-x}{\beta_r}\right)^\alpha\right], & x \geq 0 \end{cases}, \quad (4)$$

其中,

$$\beta_l = \sigma_l \sqrt{\frac{\Gamma(\frac{1}{\alpha})}{\Gamma(\frac{3}{\alpha})}}, \beta_r = \sigma_r \sqrt{\frac{\Gamma(\frac{1}{\alpha})}{\Gamma(\frac{3}{\alpha})}}, \quad (5)$$

式中:  $\Gamma(\cdot)$  为伽马函数;  $\alpha$  为控制正态分布形状的参数;  $\sigma_l^2, \sigma_r^2$  分别为 AGGD 分布的左右方差。以此, 可得到三个特征  $[\alpha, \sigma_l^2, \sigma_r^2]$ , 再通过这三个特征计算出第四个特征  $\eta$ , 计算方式为

$$\eta = (\beta_r - \beta_l) \frac{\Gamma(\frac{2}{\alpha})}{\Gamma(\frac{1}{\alpha})}. \quad (6)$$

除此之外, 计算图 2 统计量的偏度值  $D_{skew}(x)$  和峰度值  $D_{kurt}(x)$ , 组成一个特征向量  $\mathbf{F}_{CV} = [\alpha, \sigma_l^2, \sigma_r^2, \eta, D_{skew}(x), D_{kurt}(x)]$  来描述空间域的质量。

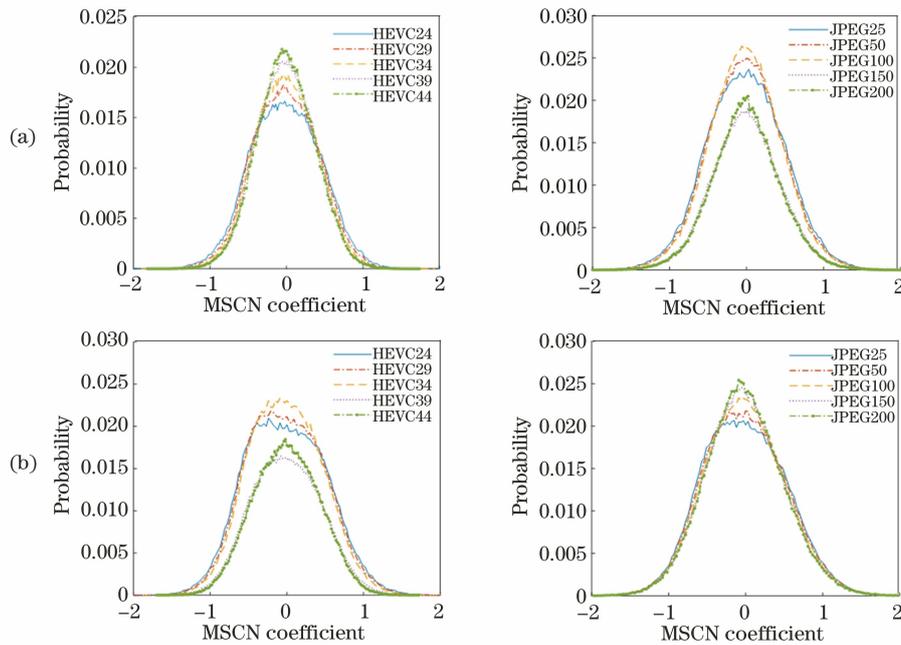


图 2 不同场景下的 MSCN 系数统计分布曲线。(a) 自行车; (b) 花

Fig. 2 Distribution curves of MSCN coefficients in different scenes. (a) Bikes; (b) Flowers

### 3.2 角度域特征

对于光场  $L(u, v, s, t)$ , 像平面上每一个点  $(s, t)$  都会接收到  $N_u \times N_v$  个不同角度方向的光线, 从而构成一个个宏像素, 用二维函数可表示为  $L_{s,t}(u, v)$ 。在光场的渲染处理中, 为了获取更多视图, 需对光场进行角度域重建, 导致角度域  $(u, v)$  的一致性发生退化, 这种退化将会明显地体现在宏像素上。

图 3 直观地显示了不同失真方式下同一宏像素的变化情况, 其中第一行 HEVC 压缩失真方式导致宏像素明显发生了模糊现象。线性插值 (LN) 和 JPEG 压缩两种失真方式虽然不会导致明显的模糊变化, 但也能看出角度一致性发生了退化, 其中各失真方式后的数字表示失真的大小, 数字越大, 表示失真越大。对此, 选取经典而易实施的 GLCM 来提取

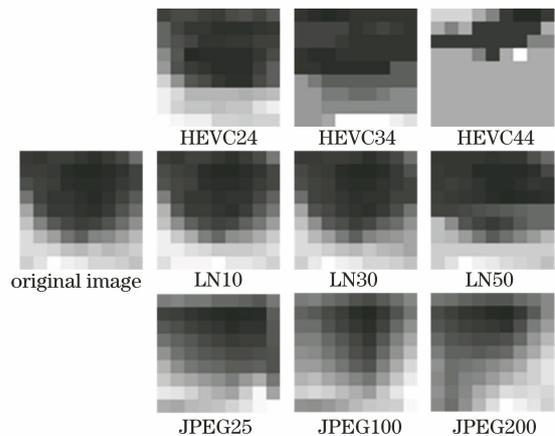


图 3 不同失真时的同一宏像素图像

Fig. 3 Macro-pixel images with different degradation level MLI 上各宏像素的特征量。

通过统计图像上保持特定距离位置上 2 个像素点灰度的联合概率分布状况, GLCM 反映图像在间隔、方向、快慢及变化幅度上的综合信息, 常用于图像特征提取<sup>[26]</sup>。一幅大小为  $M \times N$  的灰度图像  $I(x, y)$  的灰度共生矩阵可用  $P_{\Delta x, \Delta y}$  来表示:

$$P_{\Delta x, \Delta y}(i, j) = \begin{cases} 1, & I(x, y) = i \text{ and } I(x + \Delta x, y + \Delta y) = j \\ 0, & \text{otherwise} \end{cases}, \quad (7)$$

式中:  $i$  和  $j$  均为 GLCM 的大小, 计算时将图像  $I(x, y)$  的灰度划分为 8 个等级, 因此  $i = 1, 2, 3, \dots, 8, j = 1, 2, 3, \dots, 8; \Delta x$  和  $\Delta y$  为偏移量, 通过它们可以表示所统计的两点之间的距离  $d$  和角度  $\theta$ , 考虑计算效率, 本文实验选取  $d = 1, \theta = 0^\circ$ , 即偏移量  $\Delta x = 1, \Delta y = 0$ 。

使用 GLCM 描述图像灰度的变化特征时, 还需要对其进行二次计算。文献[23]中提出了 14 种由 GLCM 计算出来的参量, 但是大多数都是有相关性的, 通常只用其中的 5 个, 分别为能量(ENE)、对比度(CON)、相关性(COR)、信息熵(ENT)和同质性(HOM)。其中 ENE 是 GLCM 元素的平方和, 能够反映图像分布均匀程度和纹理粗细度; CON 反映图像的纹理沟纹深浅程度和清晰度, 值越大, 沟纹越深, 越清晰; COR 反映图像局部灰度的相关性, 值越小, 相关性越大; HOM 则反映了图像纹理局部变化的多少, 反映纹理的同质性, 值越大, 变化越少, 局部越均匀; ENT 反映图像灰度的混乱程度, 越混乱值越大, 越规则值越小。

对于 MLI, 使用(7)式计算出每一个宏像素  $L_{s,t}(u, v)$  的灰度共生矩阵  $G_{s,t}$  后, 还需对每个  $G_{s,t}$  计算二次统计量。由图 3 的第二行和第三行可以看出, 宏像素里的像素值变化没有明显的规律性。因此选取 HOM 和 ENT 这两个最能感知局部像素随机变化的统计量作为特征, 计算方式分别为

$$F_{\text{HOM}}(s, t) = \sum_i \sum_j \frac{G_{s,t}(i, j)}{1 + |i - j|}, \quad (8)$$

$$F_{\text{ENT}}(s, t) = - \sum_i \sum_j G_{s,t}(i, j) \log_2 G_{s,t}(i, j). \quad (9)$$

经过(8)、(9)式计算后, MLI 上的每个宏像素都获得了两个特征量, 一个光场数据的特征数量将会达到  $2 \times N_s \times N_t$  个, 这个数量非常大而不利于进行回归分析。为了池化特征量, 对所有宏像素的

同质性和熵进行统计, 如图 4 所示, 各失真方式后面的数字表示失真等级。图 4(a) 为所有宏像素同质性特征量的统计结果, 随着失真变大, 分布曲线向右位移; 图 4(b) 为所有宏像素熵值统计的结果, 随着失真变大, 图形向左位移。另外可发现, 分布图形状也会发生变化, 对此求取均值  $m_2(F_{\text{HOM}})$ 、 $m_2(F_{\text{ENT}})$  和偏度  $D_{\text{skew}}(F_{\text{HOM}})$ 、 $D_{\text{skew}}(F_{\text{ENT}})$  来池化特征量。最后将池化后的特征量组成一个 4 维特征向量  $\mathbf{F}_{\text{MLI}} = [m_2(F_{\text{HOM}}), m_2(F_{\text{ENT}}), D_{\text{skew}}(F_{\text{HOM}}), D_{\text{skew}}(F_{\text{ENT}})]$  来描述角度域的一致性退化。

### 3.3 空-角耦合关系特征

EPI 作为光场数据一种较为特殊的视觉形式, 通过固定一个角度坐标和一个空间坐标而得。其中水平 EPI 通过固定角度坐标  $v$  和空间坐标  $t$  所得, 表示为  $L_{v,t}(u, s)$ 。而垂直 EPI 通过固定角度坐标  $u$  和空间坐标  $s$  所得, 表示为  $L_{u,s}(v, t)$ 。由表达式  $L_{v,t}(u, s)$  和  $L_{u,s}(v, t)$  可知, EPI 既包含了空间上的一个维度信息也包含角度上的一个维度信息, 反映光场数据空-角域之间的耦合关系。

这种特殊的耦合关系使得图像具有很明显的条纹特征, 如图 5 所示。当光场数据发生失真时便会破坏这种耦合关系, 导致 EPI 上的条纹发生模糊、扭曲、错位等现象, 如图 5(b) 所示, 在 HEVC 及 JPEG 压缩失真下, EPI 的线条变得模糊, 难以区分。而角度域上的 LN 和邻近插值(NN)失真导致 EPI 上的线条发生了明显的扭曲和错位现象, 可见角度域上的插值重建操作对空-角耦合关系的破坏相比 HEVC 与 JPEG 更加严重。通过以上的分析可知, EPI 能很好地反映出空-角耦合关系的变化情况。

为了描述 EPI 上这种条形纹理的变化, 依旧使用 GLCM 来提取特征, 与宏像素不同的是, 这里选取 ENE 来描述 EPI 线条的粗细, CON 来描述条纹间的清晰度, HOM 来描述条纹的局部变化, 能衡量条纹的扭曲和错位变化。这三个统计量的计算方式为

$$F_{\text{ENE,h}}(v, t) = \sum_i \sum_j G_{v,t}(i, j)^2, \quad (10)$$

$$F_{\text{CON,h}}(v, t) = \sum_i \sum_j |i - j|^2 G_{v,t}(i, j), \quad (11)$$

$$F_{\text{HOM,h}}(v, t) = \sum_i \sum_j \frac{G_{v,t}(i, j)}{1 + |i - j|}, \quad (12)$$

式中:  $G_{v,t}(i, j)$  为水平 EPI 的 GLCM。由(7)式计算所得, 垂直 EPI 可通过相同的方式计算得到。

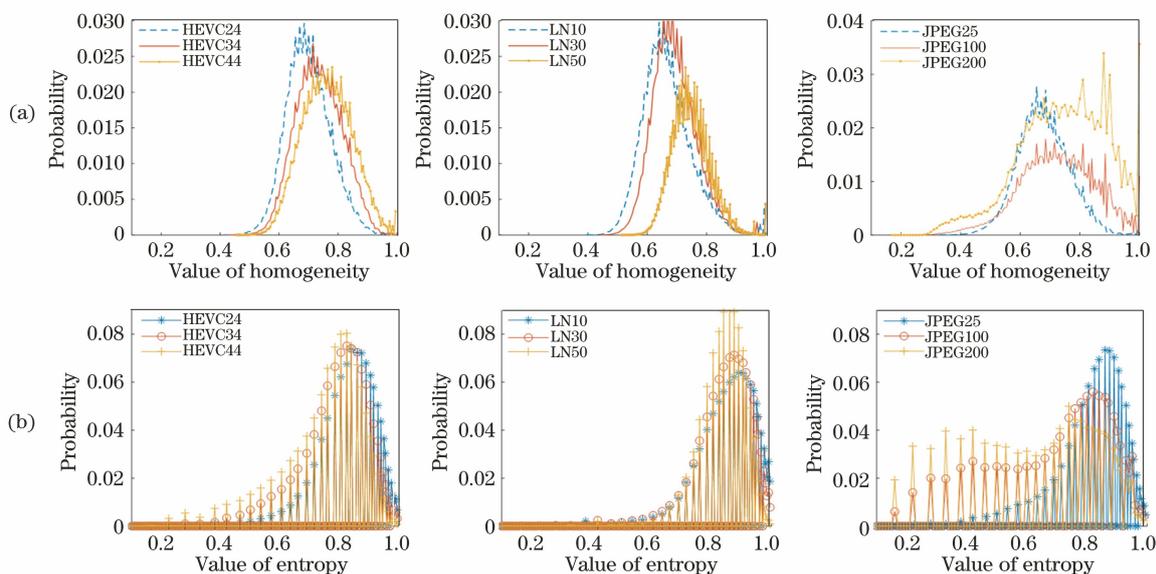


图 4 不同失真时 MLI 上两种 GLCM 特征的统计分布。(a)同质性;(b)信息熵

Fig. 4 Statistical distribution of two GLCM features on MLI with different degradation level.

(a) Homogeneity; (b) entropy

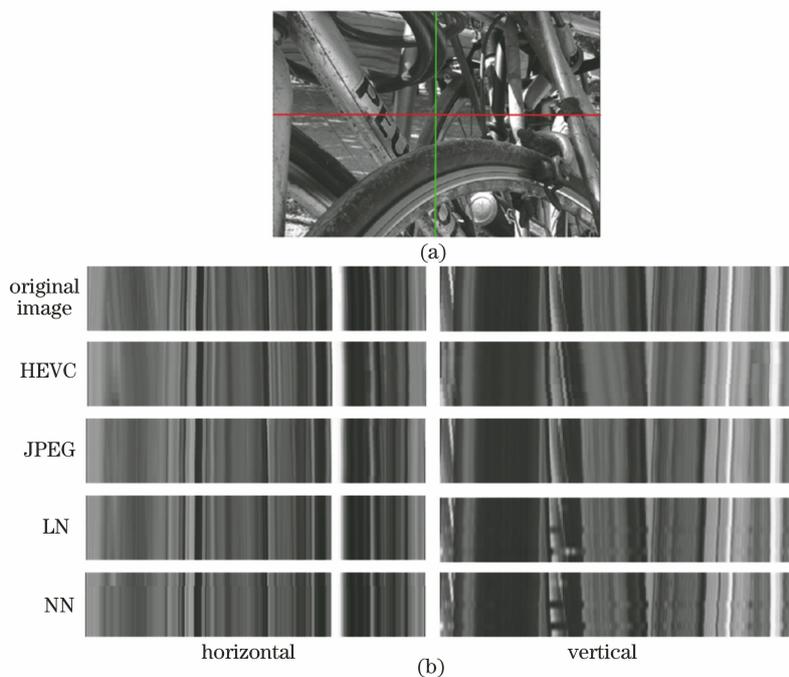


图 5 不同失真对 EPI 的影响。(a)光场中心子孔径图像;(b)不同失真时的水平和垂直 EPI

Fig. 5 Influence of different distortions on EPI. (a) Center view of light field SAI; (b) horizontal and vertical

EPI with different degradation levels

经过(10)~(12)式的计算后,从每一个 EPI 中都能提取出 3 个特征量,一个光场数据便会得到  $3 \times (N_u \times N_s + N_v \times N_t)$  个特征量。对不同失真光场数据的 3 个特征量分别进行统计,如图 6 所示,不同的失真方式导致各特征量统计分布有明显的平移现象,说明所选的特征量能较好地地区分不同失真情况。与 MLI 不同的是,EPI 数量远少于宏像素的数

量,其特征量统计分布曲线显得不平滑而导致计算偏度值不准确,因此仅对统计量进行均值池化处理,得到 6 个特征量,分别为  $f_1 = m_2(F_{ENE,h})$ ,  $f_2 = m_2(F_{CON,h})$ ,  $f_3 = m_2(F_{HOM,h})$ ,  $f_4 = m_2(F_{ENE,v})$ ,  $f_5 = m_2(F_{CON,v})$ ,  $f_6 = m_2(F_{HOM,v})$ 。这 6 个特征量再组成特征向量  $\mathbf{F}_{EPI} = [f_1, f_2, f_3, f_4, f_5, f_6]$  来描述光场空-角耦合域上的变化。

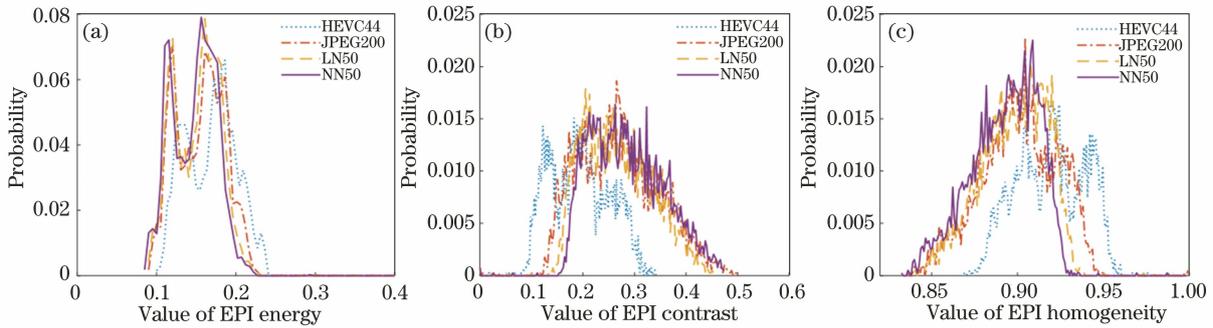


图 6 EPI 上的三种 GLCM 特征量统计分布。(a)能量;(b)对比度;(c)同质性

Fig. 6 Statistical distributions of three GLCM features on EPI. (a) Energy; (b) contrast; (c) homogeneity

### 3.4 投影域重聚焦特征

重聚焦图像<sup>[27]</sup>表示光场数据在不同深度层上的成像结果,基本原理模型如图 7 所示。穿过镜头平面和传感器平面的光线  $L_F$  向外延伸相交于重聚焦平面上,因此重聚焦面和镜头平面同样可以确定一根光线  $L_{a \cdot F}$ 。实际上  $L_F$  和  $L_{a \cdot F}$  属于同一根光线,当光线相交于重聚焦面  $a \cdot F$  的坐标  $(s, t)$  时,由相似三角形可知,其与传感器平面相交的坐标为  $(u + \frac{s-u}{a}, v + \frac{t-v}{a})$ ,所以光线  $L_{a \cdot F}$  可由  $L_F$  表示为

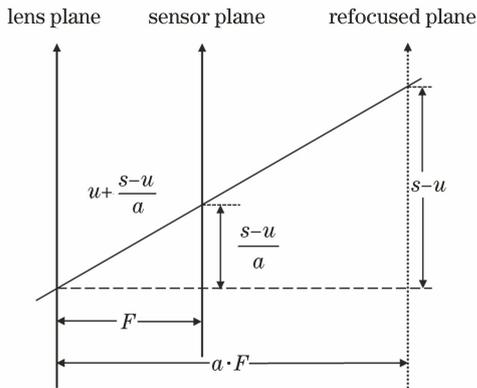


图 7 重聚焦原理模型

Fig. 7 Refocusing principle model

$$L_{a \cdot F}(u, v, s, t) = L_F \left[ u, v, u \left( 1 - \frac{1}{a} \right) + \frac{s}{a}, v \left( 1 - \frac{1}{a} \right) + \frac{t}{a} \right], \quad (13)$$

式中: $a$  为缩放因子。对(13)式进行积分,便可获得重聚焦图:

$$E_{a \cdot F}(s, t) = \iint L_{a \cdot F} \left[ u, v, u \left( 1 - \frac{1}{a} \right) + \frac{s}{a}, v \left( 1 - \frac{1}{a} \right) + \frac{t}{a} \right] dudv. \quad (14)$$

通过(14)式可知,选择不同的  $a$  便可得到不同深度层上的重聚焦图像  $E_{a \cdot F}(s, t)$ ,取  $a = \{0.6, 0.8, 1, 1.2, 1.4\}$ ,以此获得 5 个深度层的重聚焦图。图 8 为  $a = \{0.8, 1, 1.2\}$  时重聚焦面上的图像,可见清晰部位由近景推及远景,而其他部分则处于模糊失真的状况。

当光场数据发生失真时,重聚焦图像失去纹理特性而趋向平滑,导致信息熵减小。根据这一特点,对每个重聚焦图进行分块求熵,分块大小为  $8 \times 8$ ,同时为了不丢失重要的纹理信息,块与块之间步长设为 4,每一块图像熵的计算为

$$E_{\text{patch}} = - \sum_n P_n \log_2 P_n, \quad (15)$$

式中: $P_n$  为图像块中第  $n$  个灰度值出现的概率。

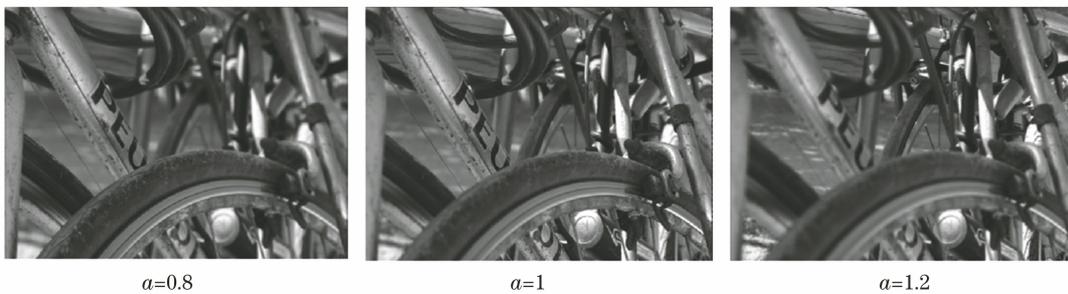


图 8 不同深度层上的重聚焦图

Fig. 8 Refocused images on different depth layers

通过分块求出的特征量数量较多,因此需要进行池化处理,再对所有块的熵进行统计。图 9 为不同失真光场在重聚焦位置  $a=1$  时的分块熵值统计分布,可见不同失真下的统计分布曲线具有不同的均值和偏度。由此,可以通过均值和偏度来池化每一张重聚焦图的统计量,得到一个 10 维的特征向量  $F_{RI} = \{m_2(I_k), D_{skew}(I_k)\}$  来描述重聚焦图的失真,其中  $I_k$  为第  $k$  个重聚焦图上的分块熵图,  $k = \{1, 2, 3, 4, 5\}$  对应  $a = \{0.6, 0.8, 1, 1.2, 1.4\}$  时的 5 个深度层。

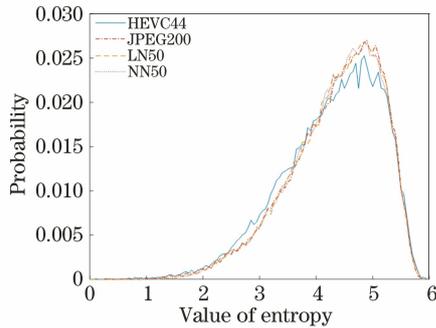


图 9 聚焦位置  $a=1$  时重聚焦图分块熵值统计分布曲线  
Fig. 9 Statistical distribution of block entropy of refocused image at focus position  $a=1$

## 4 实验与结果分析

### 4.1 光场评价数据集和性能评价指标

为了能评价光场的整体质量,先将提取的各部分特征量聚合成特征向量  $F = \{F_{CV}, F_{MLI}, F_{EPI}, F_{RI}\}$ ,然后使用 SVR 训练一个评分模型来映射此特征量对应的分数。评分模型训练实验中分别使用 SMART<sup>[9]</sup> 和 Win5-LID<sup>[11]</sup> 两个公开数据集,这两个光场数据集中的数据都由微透镜阵列相机拍摄所得,包含多种失真类型。

SMART 数据集包含由 Illum 光场相机拍摄的 16 个光场数据,每个光场数据具有  $15 \times 15$  的角度分辨率和  $434 \times 625$  的空间分辨率。同时每个光场数据引入常见的 4 种失真类型: JPEG 压缩 (JPEG)、JPEG2000 (J2K)、HEVC 帧内编码压缩、稀疏集和视差编码 (SSDC)<sup>[28]</sup>,且每个失真类型具有 4 个失真等级,总共具有 256 张失真光场数据。

Win5-LID 数据集包含 6 个由 Illum 光场相机拍摄的真实场景光场数据和 4 个模拟光场数据,真实场景光场数据的角度分辨率为  $9 \times 9$ ,空间分辨率为  $434 \times 625$ ,模拟光场数据的角度分辨率与真实场景一样,但空间分辨率为  $512 \times 512$ 。每个光场数据

都引入了 HEVC、JPEG2000、LN、NN 及两种 CNN 模型失真,除了两种 CNN 模型失真,其他每种失真包括 5 个强度等级,总共有 220 个失真光场数据。

为了将每个客观分数映射到与主观分数相同的尺度上,采用 5 参数的 Logistic 非线性回归函数<sup>[29]</sup> 进行映射,具体计算公式为

$$Q = B_1 \left\{ \frac{1}{2} - \frac{1}{1 + \exp[B_2(x - B_3)]} \right\} + B_4 x + B_5, \quad (16)$$

式中:  $x$  为客观质量评价分数;  $B_1, B_2, B_3, B_4, B_5$  均为进行拟合优化的参数。经过(16)式映射后,选择 4 个常用的评价指标来衡量客观质量评价模型的性能,即 Pearson 线性相关系数 (PLCC)、Spearman 等级相关系数 (SROCC)、Kendall 秩相关系数 (KROCC) 和均方根误差 (RMSE)。其中, PLCC、SROCC 和 KROCC 值越高,意味着与主观评分的一致性越好,说明算法性能越好。RMSE 值越小,表示算法性能越好。

### 4.2 结果分析

在训练回归预测模型时,对数据集中提取出的特征量进行归一化处理,然后从中随机选取 80% 作为训练集,剩下的 20% 作为测试集进行实验。为了保证所得指标的可靠性,进行了 1000 次计算,选取每个指标的中值作为最终的性能指标。图 10 显示了在两个数据集上运行 1000 次后各指标的分布情况,各指标的盒型形状较扁平,大部分数据聚集于中值附近,反映此模型所得分数具有较好的稳定性,提取的特征量有效。

为了进一步验证所提方法的性能,除了与几个经典的全参考二维图像质量评价方法 PSNR、SSIM<sup>[18]</sup>、MS-SSIM<sup>[30]</sup>、FSIM<sup>[31]</sup>、VSI<sup>[32]</sup>、VIF<sup>[33]</sup> 和 VSNR<sup>[34]</sup> 进行对比之外,还与近年提出的光场全参考评价方法 MDFM<sup>[14]</sup> 和 LFG-LFC<sup>[35]</sup>、光场半参考评价方法 LF-IQM<sup>[17]</sup>、光场无参考评价方法 LF-QMLI<sup>[19]</sup> 进行了对比分析。表 1、2 为两个数据集上的方法对比分析,表 1 中的对比数据来源于文献 [35],表 2 中的对比数据来源于文献 [19]。由于在 Win5-LID 数据集上的实验未计算 KROCC 这个指标,所以在表 2 中不考虑这一指标的对比。另外表 1、2 中的对比方法不一样多,是因为有些方法没有相关数据集上的实验数据。

SMART 数据集上的对比结果如表 1 所示。可以看出:所提方法的所有性能指标都比其他方法好;一些主流的图像全参考评价方法具有一定效果,但

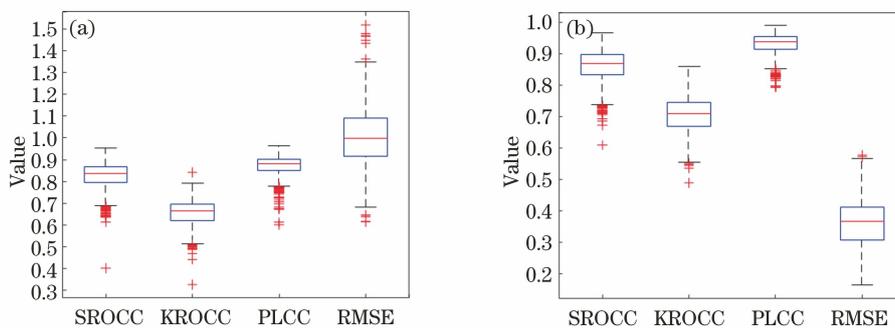


图 10 1000 次测试实验各指标分布的盒型图。(a) SMART 数据集上的结果;(b) Win5-LID 数据集上的结果

Fig. 10 Box plots of each index distribution in 1000 trials. (a) Results on SMART dataset;  
(b) results on Win5-LID dataset

对于高维光场图像的评价略显不足,主要是因为传统的图像评价方法只考虑了光场子孔径图像的失真,忽略了其他视觉形式的特征信息,导致性能指标分数较低;MDFM 方法是一种针对密集型光场数据的全参考评价方法,同样只考虑了空间域上子孔径图像的信息,仅比大多数传统方法好;LFG-LFC 除了考虑了 SAI 上的特征信息还考虑了 EPI 上的特征信息,相比其他方法,性能提高了许多;LF-IQM 是一种半参考光场评价方法,仅考虑了光场的深度信息,由于基于深度信息的估计算法本身会引入误差,因此结果表现最差。由此可见,光场的质量评价需结合多种视觉特征,这也是所提方法性能较好的原因。

表 1 在 SMART 数据集上的性能对比

Method	PLCC	SROCC	KROCC	RMSE
PSNR	0.6514	0.6883	0.4839	1.6366
SSIM	0.7180	0.6492	0.4623	1.5013
MS-SSIM	0.7578	0.6927	0.5498	1.4178
FSIM	0.7007	0.7398	0.5361	1.5389
VSI	0.7687	0.7726	0.5661	1.3796
VIF	0.5496	0.5078	0.3417	1.8020
VSNR	0.3659	0.4248	0.2798	2.0076
MDFM	0.7392	0.6468	0.4515	1.4528
LF-IQM	0.2998	0.1222	-	2.0579
LFG-LFC	0.8276	0.8246	0.6258	1.2108
Proposed method	0.8901	0.8711	0.7074	0.9246

由于 SMART 数据集只引入了 4 种较为常见的失真,为了检验所提方法是否能应对更多的失真情况,在 Win5-LID 数据集上也进行同样的对比分析。从表 2 可以看出,所提方法仅在 SROCC 这个指标上不是最优的;LF-QMLI 之所以表现较好,是因为提取了 SAI 上的高维特征和 MLI 上的熵值统

计分布特征,同时考虑了空间域和角度域上的失真。另外,通过表 1、2 发现,相比其他指标,RMSE 这个指标在两个数据集中的差别较大,主要是因为 SMART 数据集中的光场数据没有去掉各微透镜之间的黑色像素,而 Win5-LID 数据集中去掉了这些黑色像素。这些黑像素的存在会导致在计算特征时出现较多异常值;又由于 RMSE 计算每个预测值和主观值差异的累加,而 SROCC、KROCC 和 PLCC 计算的是预测值与主观值之间的相关性,对于异常值没那么敏感。所以同一方法的 RMSE 在两个数据集中的差异大。

表 2 在 Win5-LID 数据集上的性能对比

Table 2 Performance comparison on Win5-LID Dataset

Method	PLCC	SROCC	RMSE
PSNR	0.6189	0.6026	0.8031
SSIM	0.7596	0.7346	0.6650
MS-SSIM	0.8388	0.8266	0.5566
FSIM	0.8318	0.8233	0.5675
VIF	0.7032	0.6665	0.7270
VSNR	0.5050	0.3961	0.8826
LF-IQM	0.4763	0.4503	0.8991
LF-QMLI	0.9038	0.8802	0.4147
Proposed method	0.9389	0.8698	0.3655

除此之外,为了分析各部分特征对质量评测的影响,在两个数据集上分别对各部分特征量单独进行实验,结果如表 3 所示。在 SMART 数据集上,EPI 的特征量  $|F_{EPI}|$  表现最好;而在 Win5-LID 数据集中,则是重聚焦图特征量  $F_{RI}$  表现最好。这主要由于 SMART 数据集的每个光场数据包含了所有视点图像,Win5-LID 却只是包含了中间部分视点图像而去掉了那些边缘视点图像。边缘视点图像大多数是较暗或者全黑的图像,对 RI 和 MLI 的影响较大,导致在 SMART 数据集上这两部分特征表现较差。EPI 上的特征在两个数据集上都有较好表

现,说明光场的空-角耦合关系能较好地反映出光场的失真变化。另外,虽然 4 种特征向量进行单独实验时得分都不高,但是合并之后却要高于任何单一

特征向量,说明各部分的特征具有互补性,同时也反映了多视觉特征融合的光场质量评价方法的优越性。

表 3 不同部分特征单独在数据集 SMART 和 Win5-LID 上的性能对比

Table 3 Performance comparison of different features on SMART and Win5-LID datasets

Index	SMART				Win5-LID			
	$F_{CV}$	$F_{MLI}$	$F_{EPI}$	$F_{RI}$	$F_{CV}$	$F_{MLI}$	$F_{EPI}$	$F_{RI}$
SROCC	0.6487	0.3596	0.7252	0.4913	0.7246	0.7776	0.7878	0.8189
KROCC	0.4766	0.2474	0.5415	0.3439	0.5522	0.5950	0.6040	0.6382
PLCC	0.7091	0.4495	0.7526	0.6712	0.7678	0.8008	0.8643	0.9030
RMSE	1.4633	1.8670	1.3958	1.5486	0.6727	0.6245	0.5254	0.4465

## 5 结 论

从光场数据的多种视觉形式出发,研究了光场数据在空间域、角度域、空-角耦合域和投影域上的失真情况,提出了一种基于多视觉特征聚合的光场质量评价方法。空间域上,通过中心子孔径图的 NSS 分布特征来表征空间清晰度失真情况;角度域上,通过宏像素上的 GLCM 特征来表征角度一致性的变化;空-角耦合域上,通过 EPI 上的 GLCM 特征来表征空-角耦合关系的变化;投影域上,通过重聚焦图上的局部熵统计特征来表征各聚焦层的失真情况。最后聚合所有特征,利用 SVR 训练评价模型,形成了对光场质量的综合评价。实验结果显示,所提方法所得分数与主观评价分数有较好的一致性。

所提方法对多视觉特征进行聚合后回归,全面提取了光场数据的视觉特征,具有运算简单、性能较好的优点。然而,在提取 MLI 上的特征时,所提方法需计算每个宏像素的特征,导致计算效率较低,对于纹理不丰富的光场,得到的效果较差,如 SMART 数据集上的结果。因此,提出一种适应于所有数据集和实测数据的无参考质量评价方案仍然是一项具有挑战的工作。基于各视觉形式上的灰度特征,进一步研究光场失真在各视觉形式上对颜色信息的影响,将有助于提出更具有普适性的评价方法。此外,使用的 SMART 和 Win5-LID 两个公开光场质量评价数据集是微透镜阵列型光场相机拍摄的光场数据,所提方法对其他类型光场数据的效果还需要进一步实验论证。所提方法对搭建类型多样化、结构完整的光场质量评价数据集具有指导意义,同时可扩展面向多类型光场的质量评价模型和体系。

## 参 考 文 献

- [1] Wu G C, Masia B, Jarabo A, et al. Light field image processing: an overview[J]. IEEE Journal of Selected Topics in Signal Processing, 2017, 11(7): 926-954.
- [2] Wang Y Q, Yang J G, Guo Y L, et al. Selective light field refocusing for camera arrays using bokeh rendering and superresolution [J]. IEEE Signal Processing Letters, 2019, 26(1): 204-208.
- [3] Fang L, Dai Q H. Computational light field imaging [J]. Acta Optica Sinica, 2020, 40(1): 0111001. 方璐, 戴琼海. 计算光场成像[J]. 光学学报, 2020, 40(1): 0111001.
- [4] Masia B, Wetzstein G, Didyk P, et al. A survey on computational displays: pushing the boundaries of optics, computation, and perception [J]. Computers & Graphics, 2013, 37(8): 1012-1038.
- [5] Narain R, Albert R A, Bulbul A, et al. Optimal presentation of imagery with focus cues on multi-plane displays[J]. ACM Transactions on Graphics, 2015, 34(4): 1-12.
- [6] Maimone A, Wetzstein G, Hirsch M, et al. Focus 3D: compressive accommodation display [J]. ACM Transactions on Graphics, 2013, 32(5): 1-13.
- [7] Levoy M, Hanrahan P. Light field rendering[C]// Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques-SIGGRAPH'96, August 1, 1996, New York, NY, USA. New York: ACM Press, 1996: 31-42.
- [8] Chang C C, Lin C J. Libsvm: a library for support vector machines[J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2(3): 1-27.
- [9] Paudyal P, Battisti F, Sjöström M, et al. Towards the perceptual quality evaluation of compressed light field images[J]. IEEE Transactions on Broadcasting, 2017, 63(3): 507-522.
- [10] Adhikarla V K, Vinkler M, Sumin D, et al. Towards a quality metric for dense light fields[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 3720-3729.
- [11] 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), October 7-10, 2018, Athens, Greece. New York: IEEE Press, 2018: 41-45.
- [12] Viola I, Ebrahimi T. VALID: visual quality assessment for light field images dataset[C]//2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX), May 29-June 1, 2018, Cagliari, Italy. New York: IEEE Press, 2018.
- [13] Amirpour H, Pinheiro A M G, Pereira M, et al. Reliability of the most common objective metrics for light field quality assessment[C]//ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 12-17, 2019, Brighton, UK. New York: IEEE Press, 2019: 2402-2406.
- [14] Tian Y, Zeng H Q, Xing L, et al. A multi-order derivative feature-based quality assessment model for light field image [J]. *Journal of Visual Communication and Image Representation*, 2018, 57: 212-217.
- [15] Fang Y, Wei K, Hou J, et al. Light field image quality assessment by local and global features of epipolar plane image [C] // 2018 IEEE Fourth International Conference on Multimedia Big Data (BigMM), September 13-16, 2018, Xi'an, China. New York: IEEE Press, 2018.
- [16] Min X K, Zhou J T, Zhai G T, et al. A metric for light field reconstruction, compression, and display quality evaluation[J]. *IEEE Transactions on Image Processing*, 2020, 29: 3790-3804.
- [17] Paudyal P, Battisti F, Carli M. Reduced reference quality assessment of light field images[J]. *IEEE Transactions on Broadcasting*, 2019, 65(1): 152-165.
- [18] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity [J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.
- [19] 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), November 12-15, 2019, Ningbo, China. New York: IEEE Press, 2019.
- [20] 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.
- [21] Ruderman D L. The statistics of natural images[J]. *Network: Computation in Neural Systems*, 1994, 5(4): 517-548.
- [22] Huang J G, Mumford D. Statistics of natural images and models[C]//Proceedings of 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 23-25, 1999, Fort Collins, CO, USA. New York: IEEE Press, 1999: 541-547.
- [23] Haralick R M, Shanmugam K, Dinstein I. Textural features for image classification [J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1973, 3(6): 610-621.
- [24] Yu W, Xu J J, Liu Y Y, et al. No-reference quality evaluation for gamut mapping images based on natural scene statistics[J]. *Laser & Optoelectronics Progress*, 2020, 57(14): 141006.  
余伟, 徐晶晶, 刘玉英, 等. 基于自然场景统计的色域映射图像无参考质量评价[J]. *激光与光电子学进展*, 2020, 57(14): 141006.
- [25] Zhao S Y, Chen Z B, Yang K, et al. Light field image coding with hybrid scan order[C]//2016 Visual Communications and Image Processing (VCIP), November 27-30, 2016, Chengdu, China. New York: IEEE Press, 2016: 1-4.
- [26] Chu L, Liu B, Xu L, et al. Speckle quality evaluation based on gray level co-occurrence matrix [J]. *Laser & Optoelectronics Progress*, 2021, 58(4): 0410019.  
初录, 刘斌, 许亮, 等. 基于灰度共生矩阵的散斑质量评价[J]. *激光与光电子学进展*, 2021, 58(4): 0410019.
- [27] Ng R. Digital light field photography[M]. Stanford: Stanford university, 2006.
- [28] Li Y, Sjöström M, Olsson R, et al. Scalable coding of plenoptic images by using a sparse set and disparities [J]. *IEEE Transactions on Image Processing*, 2016, 25(1): 80-91.
- [29] Sheikh H R, Sabir M F, Bovik A C. A statistical evaluation of recent full reference image quality assessment algorithms [J]. *IEEE Transactions on Image Processing*, 2006, 15(11): 3440-3451.
- [30] Wang Z, Simoncelli E P, Bovik A C. Multiscale structural similarity for image quality assessment[C]// The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003, November 9-12, 2003, Pacific Grove, CA, USA. New York: IEEE Press, 2003: 1398-1402.
- [31] Zhang L, Zhang L, Mou X Q, et al. FSIM: a feature similarity index for image quality assessment [J]. *IEEE Transactions on Image Processing*, 2011, 20(8): 2378-2386.
- [32] Zhang L, Shen Y, Li H Y. VSI: a visual saliency-induced index for perceptual image quality assessment [J]. *IEEE Transactions on Image Processing*, 2014,

- 23(10): 4270-4281.
- [33] Sheikh H R, Bovik A C. Image information and visual quality [J]. IEEE Transactions on Image Processing, 2006, 15(2): 430-444.
- [34] Chandler D M, Hemami S S. VSNR: a wavelet-based visual signal-to-noise ratio for natural images [J]. IEEE Transactions on Image Processing, 2007, 16(9): 2284-2298.
- [35] Tian Y, Zeng H Q, Hou J H, et al. Light field image quality assessment via the light field coherence [J]. IEEE Transactions on Image Processing, 2020, 29: 7945-7956.