

引用格式: HUANG Zefeng, YANG Shen, DENG Huiping, et al. Light Field All-in-focus Image Fusion Based on MDLatLRR and KPCA[J]. Acta Photonica Sinica, 2023, 52(4):0410004

黄泽丰,杨莘,邓慧萍,等. 基于 MDLatLRR 和 KPCA 的光场图像全聚焦融合[J]. 光子学报, 2023, 52(4):0410004

基于 MDLatLRR 和 KPCA 的光场图像全聚焦融合

黄泽丰, 杨莘, 邓慧萍, 李青松

(武汉科技大学 信息科学与工程学院, 武汉 430081)

摘要: 为了提升光场成像的空间分辨率, 结合光场图像数字重聚焦与多聚焦图像融合, 提出了一种基于多尺度潜在低秩分解和核主成分分析的光场图像全聚焦融合算法。首先, 对光场图像进行数字重聚焦得到重聚焦图像, 然后对各重聚焦图像进行多尺度分解提取出基础层和显著层, 对基础层、显著层分别采用局部梯度差值加权算法和多尺度梯度域显著性提取算法计算相应的特征系数; 其次, 联立基础层和各显著层的特征系数矩阵, 然后用核主成分分析进行降维融合得到融合特征系数矩阵, 使得经融合特征系数生成的聚焦决策图能充分考虑基础层和显著层的特征信息; 最后, 用聚焦决策图引导重聚焦图像进行全聚焦融合。实验结果表明, 该算法与传统方法相比在视觉效果和边缘信息丰富度上具有更优表现, 所生成的光场全聚焦图像具有更高的分辨率和更好的视觉效果。

关键词: 光场; 全聚焦图像融合; 数字重聚焦; 多尺度潜在低秩分解; 核主成分分析

中图分类号: TN911.73

文献标识码: A

doi:10.3788/gzxb20235204.0410004

0 引言

传统相机在拍照时会丢失光线分布的信息, 而光场相机可以同时捕获光线的角度和空间信息, 能在不减少光圈的情况下扩展相机的景深, 从而缩短曝光时间并降低图像噪声^[1]。光场相机可以在一次拍摄中直接捕获光线场, 但会牺牲成像的空间分辨率和角度分辨率, 因此所成像的空间分辨率低于原生图像传感器生成的图像^[2]。为了弥补上述缺点, 对光场进行重聚焦从而有效提升光场图像的空间分辨率。光场数字重聚焦通过将光线从主透镜平面重新投影到新的平面上来生成重聚焦图像, 经过重聚焦后的图像其聚焦区域的清晰度得到提高。多聚焦图像融合通过从同一场景多幅部分聚焦的图像中创建一个全局聚焦的图像来扩展光学透镜视场深度^[3]。光场图像全聚焦融合通过结合多聚焦图像融合与光场数字重聚焦, 从而获得全局聚焦的光场图像。传统的多聚焦图像融合算法主要分为空间域和变换域两类。空间域方法基于像素强度来处理并融合图像; 变换域方法通过某种变换将源图像分解成不同的子带图像, 然后用不同的融合规则对各子带进行融合。此外, 随着神经网络技术的发展, 基于深度学习的图像融合方法也开始崭露头角。

近两年, 基于光场数字重聚焦的全聚焦图像融合方法逐渐流行。武迎春等^[4]使用边缘增强型引导滤波对经三尺度分解后的重聚焦图像进行优化并生成全聚焦图像, 该方法能补偿因光场标定误差而丢失的边缘信息并提高了全聚焦图像的清晰度。谢颖贤等^[5]使用离散小波变换对重聚焦图像进行分解, 随后对各分量采用不同融合规则来生成全聚焦图像, 该方法避免了因融合产生的块效应并优化了图像的视觉效果。苏博妮^[6]将重聚焦图像按亮度和色度通道分别进行处理来生成全聚焦图像, 该方法能有效提高光场全聚焦图像的空间分辨率。使用光场数字重聚焦生成的重聚焦图像不受由相机调焦引起的成像范围及成像角度的限制, 为全聚焦图像的生成提供了便利。

基金项目: 国家自然科学基金(No. 61702384)

第一作者: 黄泽丰, 1050374779@qq.com

通讯作者: 杨莘, yangshen@wust.edu.cn

收稿日期: 2022-07-21; 录用日期: 2022-08-30

<http://www.photon.ac.cn>

本文提出了一种基于多尺度潜在低秩分解(Multi-level image Decomposition base on Latent Low-Rank Representation, MDLatLRR)的光场全聚焦融合算法。算法将多尺度潜在低秩分解引入光场全聚焦图像融合,经分解得到的基础层保留了图像的结构信息,显著层保留了图像的边缘信息和纹理细节。对各层的特征提取算法进行改进,基础层用局部双区域窗代替八邻域窗,使计算得到的图像锐度能更加精确;显著层使用基于引导滤波的多尺度视觉显著性提取算法,增强了视觉显著性提取能力。引入核主成分分析(Kernal Principal Component Analysis, KPCA)对各层特征系数进行融合,生成由基础层图像锐度和显著层视觉显著性共同决定的聚焦决策图。

1 光场图像数字重聚焦

与传统图像不同,光场图像通常使用四维双平面参数化模型 $L(u, v, s, t)$ 来进行表示,其中, L 表示光线的强度; (u, v) 为主透镜平面,用于控制角度分辨率; (s, t) 为传感器平面,用于控制空间分辨率。通过对双平面参数进行积分,可获得聚焦平面的积分图像,表示为

$$I(x, y) = \iint L(u, v, s, t) du dv \quad (1)$$

通过对光场重新投影,使光线由传感器平面投影到重聚焦平面上,然后对其进行积分便可得到重聚焦图像,即

$$I(s', t') = \iint L\left(u, v, u + \frac{s' - u}{\omega}, v + \frac{t' - v}{\omega}\right) du dv \quad (2)$$

式中, ω 代表聚焦系数。通过调节聚焦系数,能将光场聚焦在不同平面上。光场重聚焦参数化模型如图1所示。

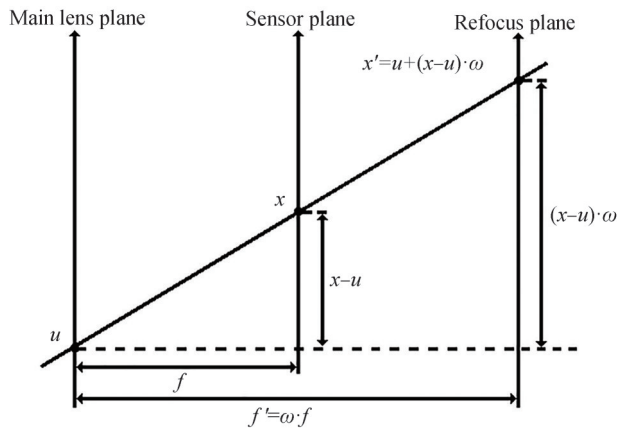


图1 光场图像数字重聚焦参数化模型

Fig. 1 Digital refocusing parameterization model of light field image

2 基于多尺度潜在低秩分解的光场图像全聚焦融合原理

基于多尺度潜在低秩分解和核主成分分析的光场图像全聚焦融合算法流程如图2所示。首先通过对4D光场图像进行光场数字重聚焦得到聚焦于不同深度的重聚焦图像。为了更为细致地提取各重聚焦图像的聚焦特征,采用多尺度潜在低秩分解将重聚焦图像分解为基础层和显著层,随后针对基础层和显著层各自的特性分别采用不同的算法进行特征提取:基础层包含重聚焦图像的结构信息,通过分析局部梯度变化情况能得到聚焦区域的大致轮廓;显著层包含重聚焦图像的纹理细节和显著目标,计算视觉显著性能反映聚焦区域的边缘信息,使得聚焦区域的判决更为准确可靠。聚焦决策图的生成由各层特征系数决定,用核主成分分析对特征系数进行降维融合,融合后的特征系数综合了基础层和显著层各自的特征。最后由融合特征系数生成聚焦决策图,引导重聚焦图像的融合得到光场全聚焦图像。

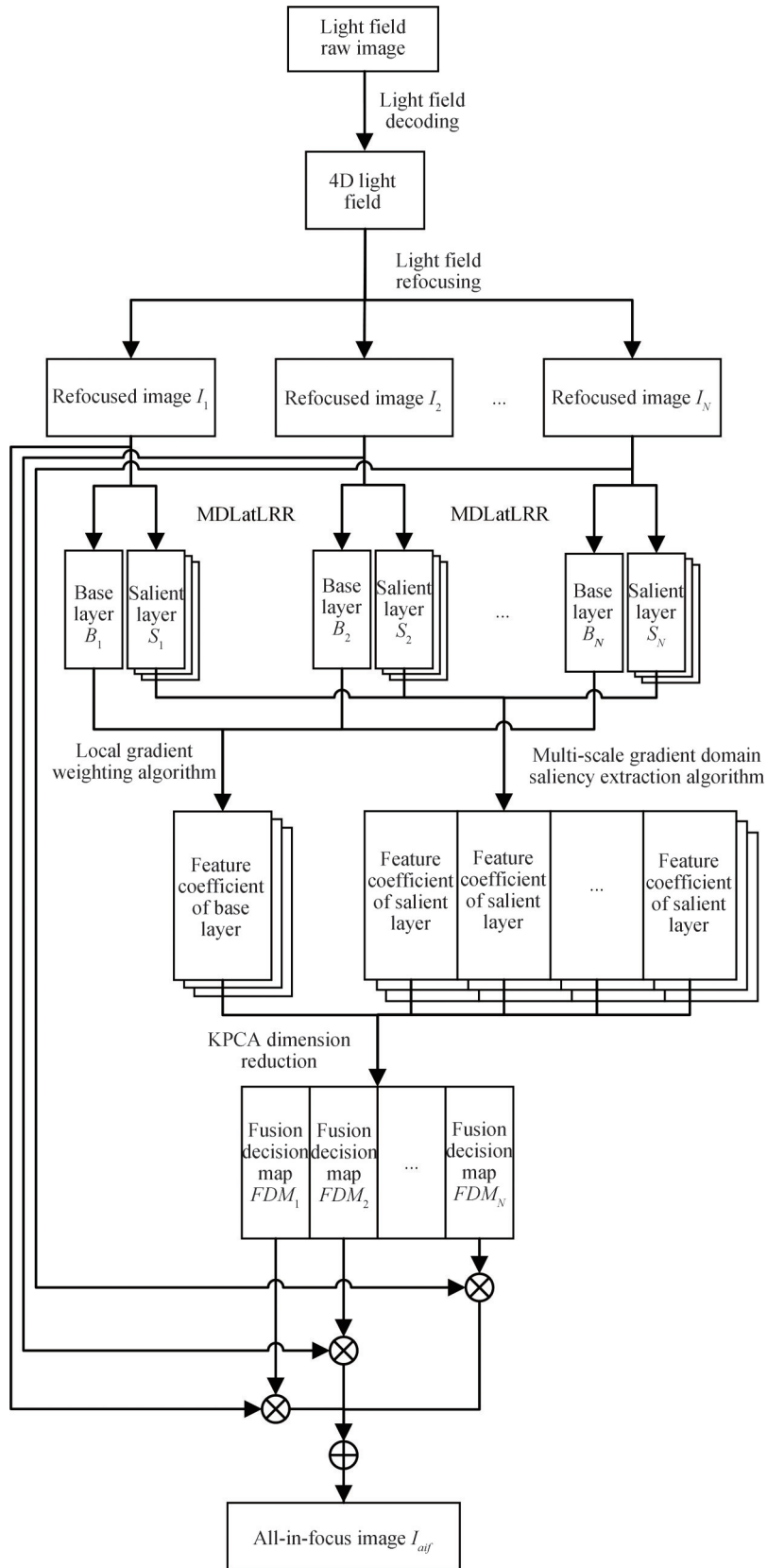


图2 基于多尺度潜在低秩分解的光场全聚焦图像融合算法流程

Fig. 2 Full-focus image fusion algorithm flow of light field based on multi-scale latent low-rank decomposition

2.1 多尺度潜在低秩分解

对于一个数据矩阵,可以将其看作在完备字典下的线性组合,具体表示为

$$X = DZ \quad (3)$$

式中, X 为数据矩阵, Z 为系数矩阵, D 为完备字典。低秩表示法在式(3)的基础上通过使用数据矩阵本身作为字典来实现数据矩阵的子空间分割。该过程可表示为

$$\min_Z \|Z\|_* \quad \text{s.t. } X = XZ \quad (4)$$

式中, $\|\cdot\|_*$ 代表矩阵的核范数, 表示矩阵的奇异值之和。低秩表示法希望系数矩阵 Z 的核范数尽可能地小。当数据采样不足时, 系数矩阵 Z 只能取单位矩阵作为唯一解从而造成低秩表示失败, 这种现象被称作隐藏效应。

潜在低秩表示法在低秩表示法的基础上将解决隐藏效应转换为对凸优化问题的求解, 具体可以表示为

$$\min_{Z, L, E} \|Z\|_* + \|H\|_* + \lambda \|E\|_1 \quad \text{s.t. } X = XZ + HX + E \quad (5)$$

式中, λ 表示大于零的平衡系数, $\|\cdot\|_*$ 和 $\|\cdot\|_1$ 分别代表核范数和 L1 范数, H 是显著系数的投影矩阵, E 是系数噪声矩阵。一旦得到投影矩阵, 就可将其用于提取图像的显著信息。

潜在低秩分解首先对投影矩阵与输入图像的乘积进行重构来得到图像的显著层, 然后使用输入图像减去显著图像得到基础层。多尺度潜在低秩分解则是对基础层继续进行潜在低秩分解, 从而得到更加丰富的显著信息。经过多尺度潜在低秩分解后, 输入图像被分解为一个基础层和若干个显著层, 基础层保留输入图像的低频信息; 显著层实现对输入图像细节纹理和显著目标等高频信息的提取; 随着迭代次数的增加, 显著层对输入图像的特征提取能力得到增强, 对细节纹理以及显著目标的提取更为精确。

本文算法使用三尺度潜在低秩分解对各重聚焦图像进行分解: 首先对输入图像进行潜在低秩分解得到初始基础层和显著层一层, 随后将初始基础层作为输入图像再次进行潜在低秩分解, 如此往复直至分解至第三层, 最后将第三次分解得到的基础层作为三尺度潜在低秩分解的基础层, 分解结果如图 3 所示。其中, 基础层包含源图像的主要结构, 如图 3(b) 所示; 显著层随着分解尺度的迭代显著细节逐渐得到突出, 如图 3(c)~(e) 所示。

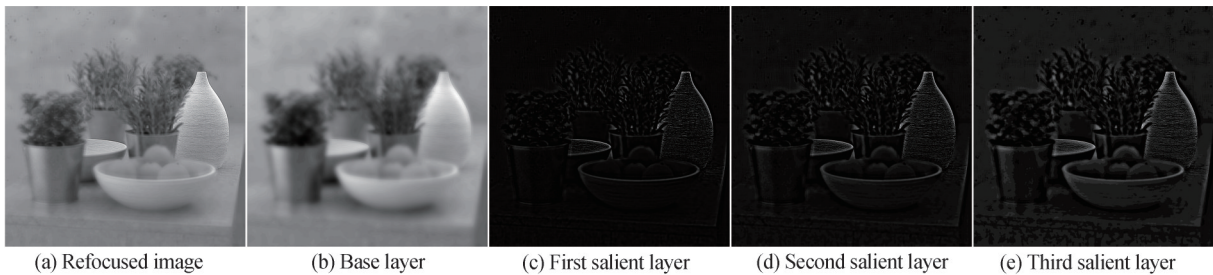


图 3 三尺度潜在低秩分解结果

Fig. 3 Results of three-scale latent low-rank decomposition

2.2 基础层特征系数矩阵求解

基础层保留了重聚焦图像中的结构信息以及因分解产生的伪影。在重聚焦图像中, 聚焦区域像素灰度值的变化通常大于散焦区域像素灰度值的变化, 而梯度反映了灰度值的变化程度。此外, 像素与其相邻像素之间的梯度差代表了图像的锐度, 计算图像的锐度可以扩大一个像素与其相邻像素之间的差异, 将有益于聚焦和散焦区域的判定。文献[7]提出了一种基于八邻域图像梯度差值的聚焦点检测算法, 该算法中单像素点图像锐度值 $S(i, j)$ 计算公式为

$$S(i, j) = \sum_{x=-1}^1 \sum_{y=-1}^1 (G(i, j) - G(i+x, j+y))^2 \quad (6)$$

式中, $G(i, j)$ 表示位于坐标 (i, j) 处像素的梯度值。该方法通过计算源图像在梯度域中每个像素与其八邻域像素的梯度差值的平方和从而得到源图像的锐度。该方法具有速度快且高效的优点, 但是在计算局部图像锐度时因易受水平和垂直方向的能量抵消而导致抗噪性能下降。

在式(6)基础上对原有方法加以改进, 将基于像素点的八邻域拓展到以像素点为中心向外辐射的 5×5 局部窗, 并根据与中心的距离将窗内像素点划分为内、外两个区域 Ω_1 和 Ω_2 , 如图 4(b) 所示。对内部区域 Ω_1

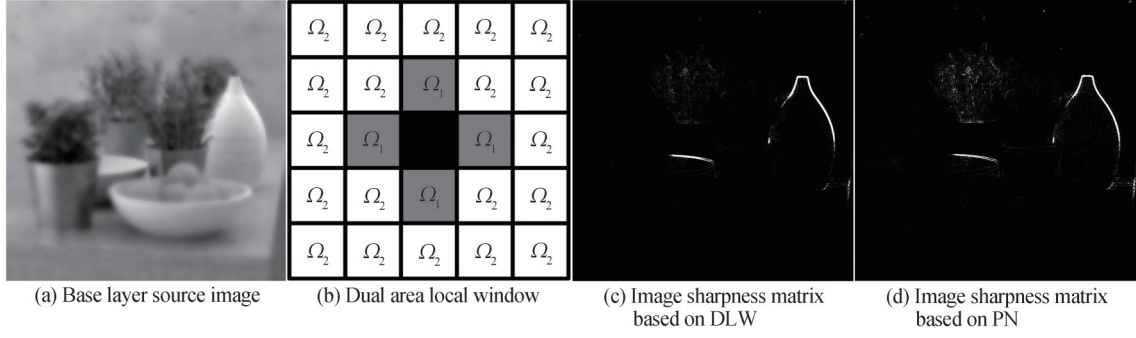


图4 局部梯度差值加权法

Fig. 4 Local step difference weighting method

使用梯度差平方和来扩大目标像素与其四邻域像素之间的差异;对外部区域 Ω_2 使用受L2范数约束的梯度差来统计周边区域对目标像素的影响。改进后的算法对噪声具有更高的鲁棒性并能保留更为丰富的梯度信息。局部锐度值 S_{local} 的具体定义为

$$S_{\text{local}}(i, j) = \frac{1}{|\Omega_1|} S_{\Omega_1}(i, j) + \frac{1}{|\Omega_2|} S_{\Omega_2}(i, j) \quad (7)$$

$$\begin{cases} S_{\Omega_1}(i, j) = \sum_{r_1 \in \Omega_1} \sum_{c_1 \in \Omega_1} (G(i, j) - G(i + r_1, j + c_1))^2 \\ S_{\Omega_2}(i, j) = \sum_{r_2 \in \Omega_2} \sum_{c_2 \in \Omega_2} \|G(i, j) - G(i + r_2, j + c_2)\|_2 \end{cases} \quad (8)$$

式中, $S_{\Omega_p}(i, j)$ 代表位于坐标 (i, j) 处的像素值在区域 Ω_p 中($p \in \{1, 2\}$)的锐度值; $|\Omega_1|$ 和 $|\Omega_2|$ 作为权重系数分别取4和20,代表对应区域包含的像素数目; $\|\cdot\|_2$ 表示L2范数。

基于本文方法和文献[7]方法的图像锐度提取结果分别如图4(c)和(d)所示。通过对比两算法生成的锐度矩阵可知:基于局部窗的图像锐度算法具有更强的鲁棒性,生成的图像锐度矩阵受散焦区域噪声的影响更小,且能更精确地反映聚焦区域边界的梯度变化情况。

2.3 显著层特征系数矩阵求解

显著层保留了源图像中的细节信息和显著特征。显著特征在重聚焦图像中表现为对聚焦区域的突出,而细节信息的保留很大程度上决定了全聚焦图像质量。通过对重聚焦图像的显著层进行视觉显著性检测可以得到较为准确的聚焦区域。MA J等^[8]提出了一种简单有效的图像视觉显著性计算方法,该算法通过衡量每个像素与图像中其他所有像素的对比度来表示该像素的显著性。对于灰度级 p ,图像中对应像素的显著性 $R(p)$ 可以表示为

$$R(p) = \eta(p) \sum_{\tau=0}^{P-1} |p - \tau| \quad (9)$$

式中, $\eta(p)$ 代表图像中灰度值 p 对应的像素数目, τ 为灰度值, P 为灰度级数。该方法虽然可以快速计算出图像的视觉显著性,但所获结果易受源图像影响进而在聚焦区域之外产生较大权值,如图5(b)中间位置处的盆景轮廓和位于右下方的果盘虚影。

针对此问题提出了一种新的多尺度图像显著性提取方法,具体步骤包括:1)将直接对源图像进行显著性矩阵计算改变为在梯度域进行;2)通过引导滤波对显著性矩阵进行优化;3)将单尺度显著性检测改进为基于引导滤波的多尺度迭代显著性检测。

首先,计算源图像中每个像素的梯度值,计算公式为

$$G(i, j) = \sqrt{\nabla^x f(i, j)^2 + \nabla^y f(i, j)^2} \quad (10)$$

用Sobel算子来计算图像梯度, $\nabla^x f$ 与 $\nabla^y f$ 分别表示源图像在水平和垂直方向上的梯度值。

其次,计算源图像在梯度域中的显著性矩阵,公式为

$$V_{n,1}^k = R(G_n^k) \quad n \in \{1, \dots, N\}, k \in \{1, 2, 3\} \quad (11)$$

式中, G_n^k 表示第 n 张重聚焦图中第 k 层显著层的梯度图像, N 为重聚焦图像总数; $R(\cdot)$ 表示图像显著性矩阵;

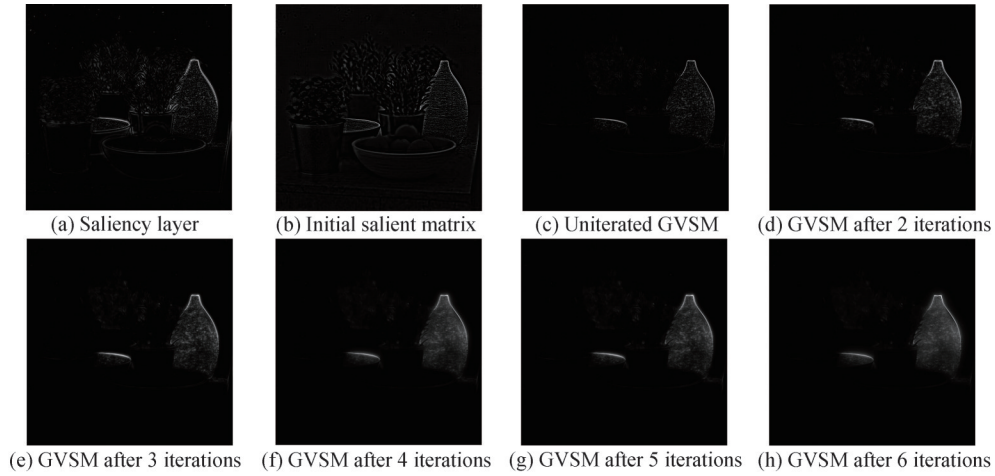


图5 多尺度迭代显著性检测算法

Fig. 5 Multi-scale iterative significance detection algorithm

$V_{n,1}^k$ 代表源图像在单尺度下的梯度域显著性矩阵。所获结果与原方法相比,位于聚焦区域之外的误判区域权值得到抑制,如图5(c)所示。

最后,将对应显著层源图像作为引导图像对单尺度梯度域显著性矩阵进行引导,随后重复该过程并使之与前次结果进行叠加,具体过程为

$$V_{n,s}^k = V_{s-1}^k + F_{\text{guidedfilter}}(f_n^k, V_{s-1}^k, s^2, \epsilon) \quad s \in \{2, \dots, M\} \quad (12)$$

式中, $F_{\text{guidedfilter}}(\cdot)$ 表示引导滤波函数, f_n^k 表示第 n 张重聚焦图中第 k 层显著层源图像, s 表示当前尺度系数, M 为总迭代次数, ϵ 为引导滤波正则化系数,本文设为0.001。在不同尺度下使用显著层源图像作为引导图像对显著性矩阵进行引导滤波,可以使显著性矩阵保留引导图像的结构且不丢失边缘信息。在不同迭代次数下生成的显著性矩阵结果如图5(d)~(h)所示。从图中可以看出,随着迭代次数的增加,聚焦区域的突出程度也越来越高,当迭代次数达到4次时,对聚焦区域的提取已相当精准,之后的迭代对于显著性提取影响甚微。

为了进一步对比不同尺度下显著层的特征提取能力,选用方差、信息熵和均方根误差等指标对1至6次迭代下梯度域显著性提取结果进行对比,其中:方差和信息熵分别描述显著性矩阵的细节保留以及显著信息提取的能力,方差越高反映了保留的细节和纹理越丰富,信息熵越高则反映显著性矩阵对输入图像的显著信息提取能力更强;此外,通过计算显著性矩阵和输入图像梯度图之间的均方根误差,能反映不同迭代次数下的显著目标提取准确度,越小的均方根误差表示检测效果越精确。为统筹考虑各项指标,对各结果经过归一化后的三个指标进行叠加得到综合评定分数,定义为

$$E_{\text{score}}(V_s) = Q_{\text{variance}}(V_s) + Q_{\text{entropy}}(V_s) + (1 - Q_{\text{rmse}}(V_s, G)) \quad (13)$$

式中, $V_s, s \in \{1, \dots, 6\}$ 表示在1至6次迭代次数下生成的显著性矩阵; Q_{variance} 、 Q_{entropy} 和 Q_{rmse} 分别计算显著性矩阵的方差、信息熵以及与梯度图像 G 的均方根误差,并在计算完后对结果进行归一化。图6反映了不同迭

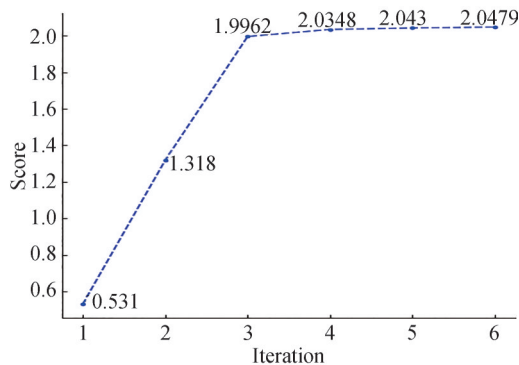


图6 不同迭代次数下显著性矩阵得分情况

Fig. 6 Score of salient matrix under different iterations

代次数下各显著性矩阵的得分情况,可以看出随着迭代次数增加,显著性矩阵的性能得到提升,并在第三次达到拐点随后增长缓慢。为了减少计算量并降低时间成本,将迭代次数选定为4次。

2.4 基于核主成分分析的聚焦决策图生成及融合

核主成分分析^[9]是主成分分析的一种非线性推广,将输入数据映射到高维特征空间中,然后对主成分进行分析。该方法可将高维数据降至低维并保留主要成分以及提取特征信息,且其效果优于同类其他方法^[10]。在本文中,光场重聚焦图像经三尺度潜在低秩分解得到了一个基础层和三个显著层,算法框架如图7所示。聚焦决策图生成详细步骤为:

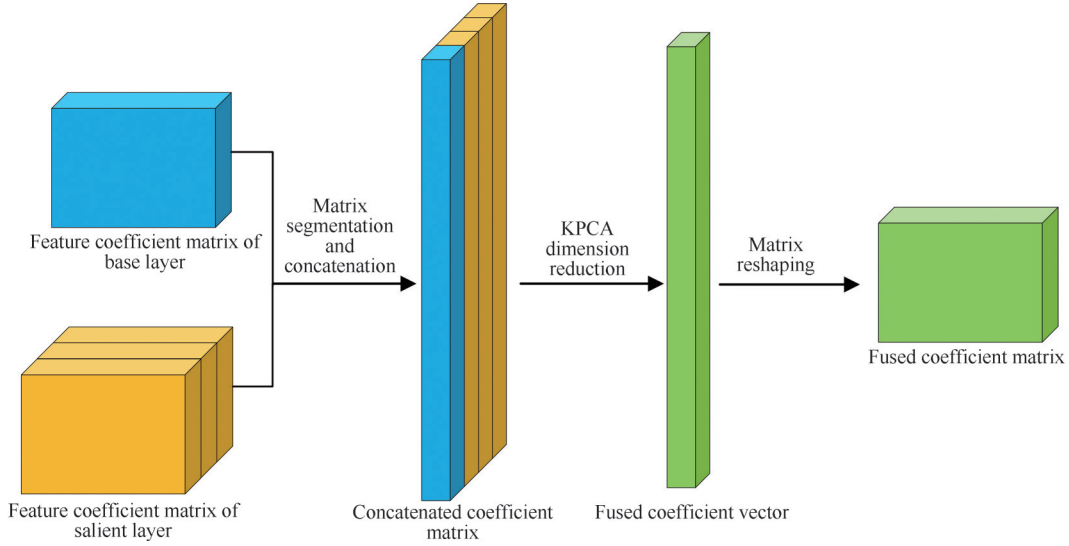


图7 基于核主成分分析的特征系数矩阵融合算法框架

Fig. 7 Framework of eigencoefficient matrix fusion algorithm based on kernel principal component analysis

1)为提高算法运行效率,对基础层和显著层的特征系数矩阵进行分割,将特征系数矩阵分解为 M^2 个子矩阵,该过程可表示为

$$F_n = \begin{bmatrix} f_{n_{11}} & \cdots & f_{n_{1M}} \\ \vdots & \ddots & \vdots \\ f_{n_{M1}} & \cdots & f_{n_{MM}} \end{bmatrix} \quad (14)$$

式中, $f_{n_{ij}}$ 表示子矩阵, i 和 j 为子矩阵在原矩阵中的对应位置。子矩阵的尺寸会影响降维的时间成本和降维融合效果,所用源图像大小为 1024×1024 ,在实验过程中分别使用 32×32 、 16×16 和 8×8 的子矩阵,经过对比发现: 32×32 子矩阵所需要的时间成本是其余尺寸的2倍及以上; 8×8 子矩阵时间成本最低,但由于尺寸过小导致最终降维融合的效果不如 16×16 子矩阵。综合考虑,将子矩阵大小定为 16×16 ,则 $M = \frac{1024}{16} = 64$ 。分割完成后,对所有子矩阵进行向量化重组,在对其他特征系数矩阵进行同样操作后将所有向量进行组合,即

$$\left\{ \begin{array}{l} F_n = \begin{pmatrix} v_{n_{11}} \\ \vdots \\ v_{n_{1M}} \\ v_{n_{21}} \\ \vdots \\ v_{n_{MM}} \end{pmatrix} \\ F_{\text{combine}} = [F_1 \quad \cdots \quad F_Y] \end{array} \right. \quad (15)$$

式中, $v_{n_{ij}}$ 是对式(14)中的子矩阵 $f_{n_{ij}}$ 进行向量化后的结果; Y 为4,表示特征系数向量总数。各分割特征系数向量经过组合后得到分割特征系数矩阵 F_{combine} ,将原本三维的数据降至二维。

2)使用核主成分分析对分割特征系数矩阵进行降维处理:先用非线性映射 $\Phi(x_i)$ 将输入的分割特征系数矩阵 F_{combine} 中的元素 $x_i(i=1, \dots, Y)$ 映射到高维特征空间中使其线性可分,然后在这个高维空间中进行主成分分析降维。高维特征空间中的协方差矩阵定义为

$$C = \frac{1}{Y} \sum_{i=1}^Y \Phi(x_i) \Phi(x_i)^T \quad (16)$$

由于非线性映射 $\Phi(x_i)$ 是未知的,需要通过选取合适的核函数 k ,将对协方差矩阵 C 的操作转化为对核矩阵 K 特征值求解的问题。核矩阵 K 表现形式为

$$K = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_Y) \\ \vdots & \ddots & \vdots \\ k(x_Y, x_1) & \cdots & k(x_Y, x_Y) \end{bmatrix} \quad (17)$$

采用高斯核函数,其定义为

$$k(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (18)$$

σ 为标准差。可知, $K_{ij} = k(x_i, x_j) = \Phi(x_i) \Phi(x_j)$ 。随后,求解核矩阵 K 的特征值 λ 以及经过标准化的特征向量 \bar{v} 。特征向量 v 的标准化定义为

$$\bar{v} = \frac{v}{\sqrt{\lambda}} \quad (19)$$

由求得的标准化特征向量 \bar{v} 以及核函数 $k(x_i, x_j)$,可以得到特征系数矩阵 F_{combine} 中的每个元素 F_i 对应的核主成分 $m(i), i \in \{1, \dots, Y\}$

$$m(i) = \sum_{j=1}^Y \bar{v}_i(j) \cdot k(x_i, x_j), \quad i = 1, \dots, Y \quad (20)$$

为了实现非线性映射的中心化,还需要对核矩阵 K 进行中心化处理。定义 O_Y 为 $Y \times Y$ 的全1矩阵,则经过中心化后的核矩阵 \bar{K} 表示为

$$\bar{K} = K - \frac{1}{Y} (K \cdot O_Y) - \frac{1}{Y} (O_Y \cdot K) + \frac{1}{Y^2} (O_Y \cdot K \cdot O_Y) \quad (21)$$

3)分割特征系数 F_{combine} 经过降维得到了融合特征系数向量,对每个重聚焦图像的融合特征系数向量进行重塑得到对应的重塑系数矩阵 $F_{\text{fused}}^n, n \in \{1, \dots, N\}, N$ 为重聚焦图像总数。采用取最大值规则对所有融合特征矩阵中的每个像素点进行判决,得到初始决策图 I_{initial} 中每个像素值的表示

$$I_{\text{initial}}(i, j) = \begin{cases} 1, & F_{\text{fused}}^n(i, j) = \max\{F_{\text{fused}}^1(i, j), \dots, F_{\text{fused}}^N(i, j)\} \\ 0, & \text{others} \end{cases} \quad (22)$$

由于初始决策图的构建可能受聚焦区域中存在的一些散焦块影响从而产生小的误判区域,所以对初始决策图执行小结构去除来优化决策图。具体过程可以表示为

$$I_{\text{final}} = F_{\text{ssr}}(I_{\text{initial}}, \psi) \quad (23)$$

式中, $F_{\text{ssr}}(\cdot)$ 为小结构去除函数, ψ 为判决阈值,当图像中存在像素数目小于 ψ 的区域时会被去除, ψ 设置为各张重聚焦图像初始决策图中聚焦区域像素总数的5%。

得到最终决策图后,使用重聚焦图像作为引导图像来对聚焦决策图进行引导滤波从而削减模糊边界的影响,全聚焦图像 I_{aif} 可以表示为

$$I_{\text{final}}^n = F_{\text{guided filter}}(I^n, I_{\text{final}}^n, r_f, \epsilon_f), \quad n \in \{1, \dots, N\} \quad (24)$$

$$I_{\text{aif}} = I_{\text{final}}^1 \cdot I^1 + \dots + I_{\text{final}}^N \cdot I^N \quad (25)$$

式中, N 代表光场重聚焦图像的数量, r_f 和 ϵ_f 分别为引导滤波函数尺度系数和正则化系数, I 代表光场重聚焦图像; I_{initial} 代表最终决策图。

3 实验结果与分析

实验使用的硬件配置为:CPU: Intel Core i7-11800H; RAM: 32.0 GB; GPU: GeForce RTX 3060, 实验环

境为 Matlab R2019a。实验分为两部分:1)选用多个4D光场子孔径图像组进行全聚焦融合实验以验证本文算法的有效性;2)与其他先进的多聚焦图像融合方法从主观和客观两方面进行比较以评估本文算法性能。

3.1 光场全聚焦融合实验

实验数据选用HCI数据集^[11]提供的4D光场数据,每组数据均包含10张重聚焦图像,实验所用部分数字重聚焦图像及对应聚焦决策图如图8所示。图8展示了从4组光场数据集中选出的3张光场数字重聚焦图像以及由本文方法生成的聚焦决策图。图9展示了4组全聚焦融合实验的总聚焦决策图以及对应的光场全聚焦图像,其中,第一、第三列为总聚焦决策图;第二、第四列为全聚焦图像。通过对比图8和图9可以发现,总聚焦决策图中不同颜色对应各单一聚焦决策图,此外,本文方法在处理聚焦区域模糊边界问题上用重聚焦图像作为引导图像对聚焦决策图进行引导滤波,使各聚焦决策图的衔接更为平滑,表现在总聚焦图上为不同颜色区域衔接处具有模糊效果。

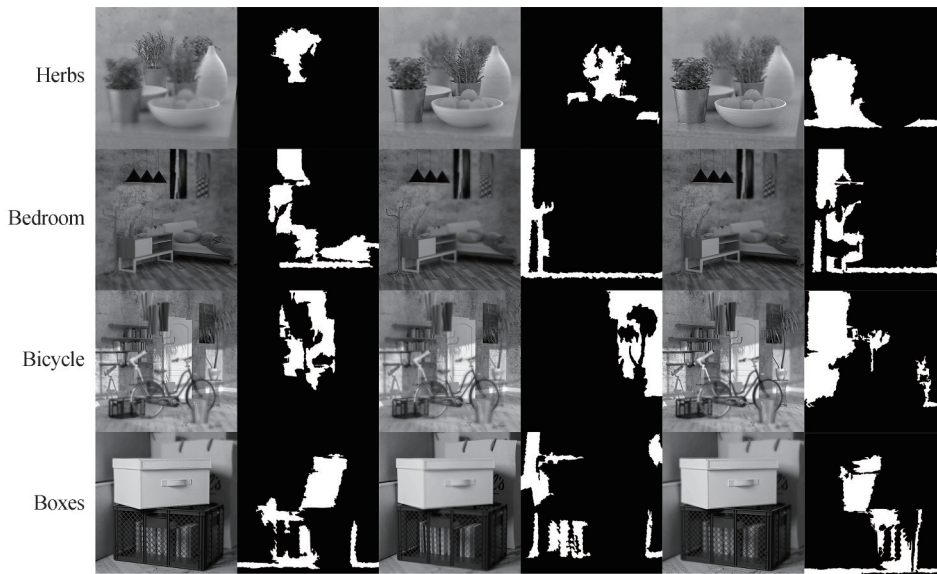


图8 4组光场重聚焦图像以及对应聚焦决策图展示

Fig. 8 Four groups of light field refocusing images and corresponding focusing decision map

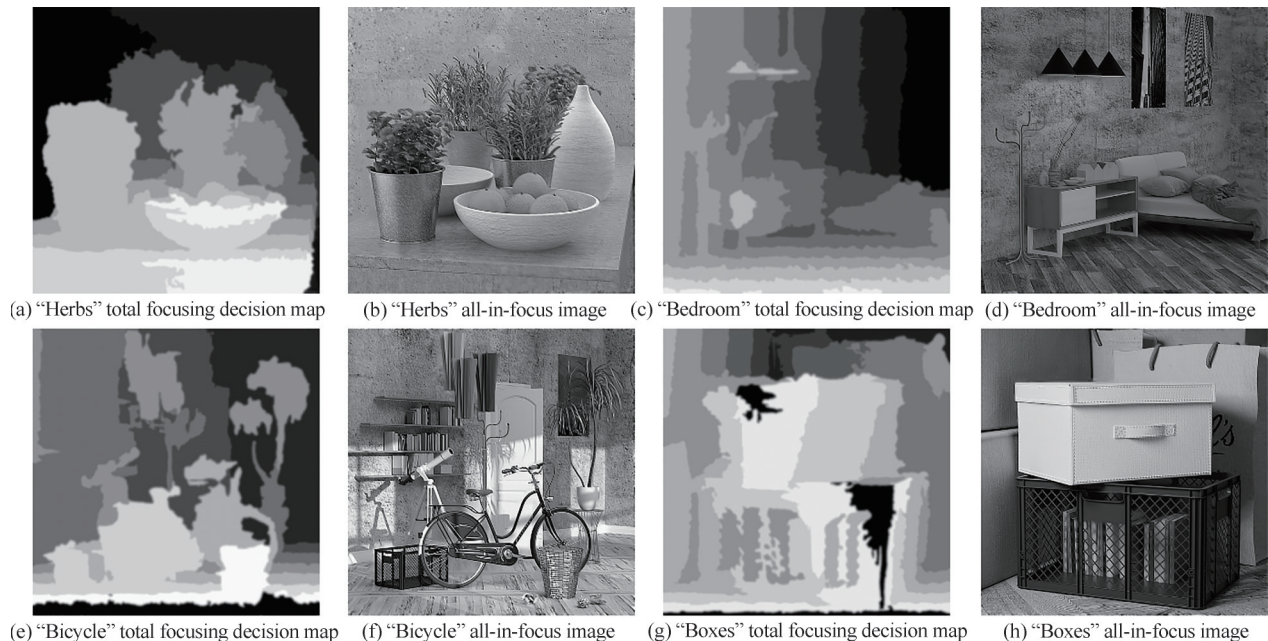


图9 各重聚焦图像总聚焦决策图以及对应全聚焦图像展示

Fig. 9 The total focusing decision map of each refocusing image and the corresponding all-in-focus image are displayed

图9(a)和(b)为Herbs图像组的总聚焦决策图和全聚焦图像,总聚焦决策图中10个灰度级对应10张重聚焦图像的聚焦区域,各聚焦区域囊括了Herbs图像的所有要素:背景(墙壁)、物品(桌子上的盆景、花瓶以及果盘)、各物品间的区域(果盘、空碟和盆景三者之间的区域)以及前景(桌角和桌子外侧)。通过总聚焦决策图引导重聚焦图像进行融合便能生成全聚焦图像,观察图9(b)、(d)、(f)及(h)等全聚焦融合结果可以发现:各聚焦区域边界衔接自然、边缘信息得到良好保留,果盘中各水果轮廓以及白色立式花瓶和其左侧的植物相交处;同时,经过重聚焦后的光场图像实现了对纹理细节的增强(植物叶子、花瓶上的花纹以及背景墙上的细节),使全聚焦融合结果具有更好的视觉效果。此外,本文算法不对分解后的各层直接进行融合,而是通过生成聚焦决策图来引导重聚焦图像进行融合,所获得的全聚焦图像不会因为各层融合产生的噪声而导致图像质量下降。

3.2 多聚焦图像融合对比实验

为了进一步评估本文算法的性能,选取了6种多聚焦图像融合算法:RW^[12]、f-PCNN^[13]、IFCNN^[14]、MGFF^[15]、MST^[16]、CSR^[17]来进行比较。对比实验采用LFSD^[18]数据集提供的光场重聚焦图像,由于大多数方法仅针对两幅重聚焦图像的融合,所以从每组重聚焦图像中选出两张作为输入图像,共计10组。为了保证实验结果客观公正,所选方法均使用原文提供的源代码以及最优参数。选取四大类共10种客观评价指标^[19]:1)基于人类视觉系统的指标,包括 Q_{cv} [20]、 Q_{cb} [21]、 Q_{abf} [22], Q_{cv} 考虑局部测量来估计源图像中重要信息被融合图像的代表程度, Q_{cb} 基于人眼视觉系统模型对融合图像进行质量评价, Q_{abf} 利用人眼对比度灵敏度函数比较融合图像与源图像的视觉差异;2)基于图像梯度的指标,包括空间频率(Spatial Frequency, SF)、平均梯度(Average Gradient, AG)以及边缘强度(Edge Intensity, EI);3)基于统计学的指标,包括方差(Variance, Var);4)基于信息论的指标,包括信息熵(Entropy, EN)、交叉熵(Cross Entropy, CE)以及互信息量(Mutual Information, MI)。这四类指标涵盖了对人眼视觉感知、图像显著性、边缘信息量以及图像失真程度等因素的考虑,能够较为全面且客观地评价融合图像的质量。

图10反映了实验中7种算法在融合LoveLocks图像组时的表现。其中,图10(a)和(b)均为重聚焦图像,(c)~(i)为所选算法各自的融合结果。为进行更细致的对比,对融合结果的细节放大处理,红框和绿框即为局部放大区域。通过对比放大区域可以发现,本文算法在对源图像的细节保留上有较为优秀的表现:右下角红框放大区域展示了各方法对锁面数字细节的保留情况,可以看出本文方法保留的数字更为清晰,每个数字都具有最好的边缘视觉效果,更符合人眼的视觉特性;右上角绿框放大区域展示了各方法对右聚焦图像与左聚焦图像模糊边界接壤处复杂纹理的融合情况,可以发现CSR方法在该区域产生了明显的伪影,IFCNN和MGFF等方法的融合结果丢失了“L”型锁的表面细节,使得锁面看起来非常平滑,本文方法较f-PCNN、MST和RW等方法具有最高的对比度,更能凸显“L”型表面凹凸不平的细节。7种方法在LoveLocks图像组上的融合结果客观指标如表1所示,其中,最优值以红色表示,次优值以蓝色表示。10种评价指标中,除 Q_{cv} 和交叉熵(CE)为越小越好外,其余均为数值越大越好。可以看出,本文方法的融合结果在10个指标中具有3个第一和4个第二,在 Q_{cv} 、互信息量以及交叉熵均取得了第一名,说明经本文算法生成的融合图像较好地保留了源图像的细节纹理并具有更优的视觉表现。

图11与图10具有相同的内容分布,反映了7种算法在融合Edelweiss图像组时的表现。红框放大部分反映了各方法对啤酒商标的细节保留情况,可以发现CSR、IFCNN和RW等方法产生了明显的伪影,MGFF方法未产生伪影但其商标清晰度有所下降,f-PCNN、MST和本文方法在主观上均有较好表现,但从表2中的数据可以看出本文方法具有最好的互信息量以及信息熵,说明本文方法对源图像的信息保留程度更高。绿框放大区域反映了在面对左右聚焦图像模糊边界区域时,各方法对背景(天空等均匀区域)和显著目标(教堂塔尖)的融合情况,可以看出,除RW和本文方法外,其余5种方法在塔尖区域均有非常明显的伪影,进一步仔细对比可以发现,经本文方法融合后的结果没有伪影,塔尖与天空融合后的视觉效果要好于RW方法。

表2反映了7种方法在Edelweiss图像组上的融合结果客观指标对比,本文方法在除 Q_{cv} 和交叉熵以外的8种指标上均取得较为优秀的表现。 Q_{cv} 反映了源图像中的重要信息在融合图像中的表示程度,由表2可知,RW方法和CSR方法的Edelweiss融合结果在 Q_{cv} 指标上取得了良好的成绩,但在图11中的红框放大区

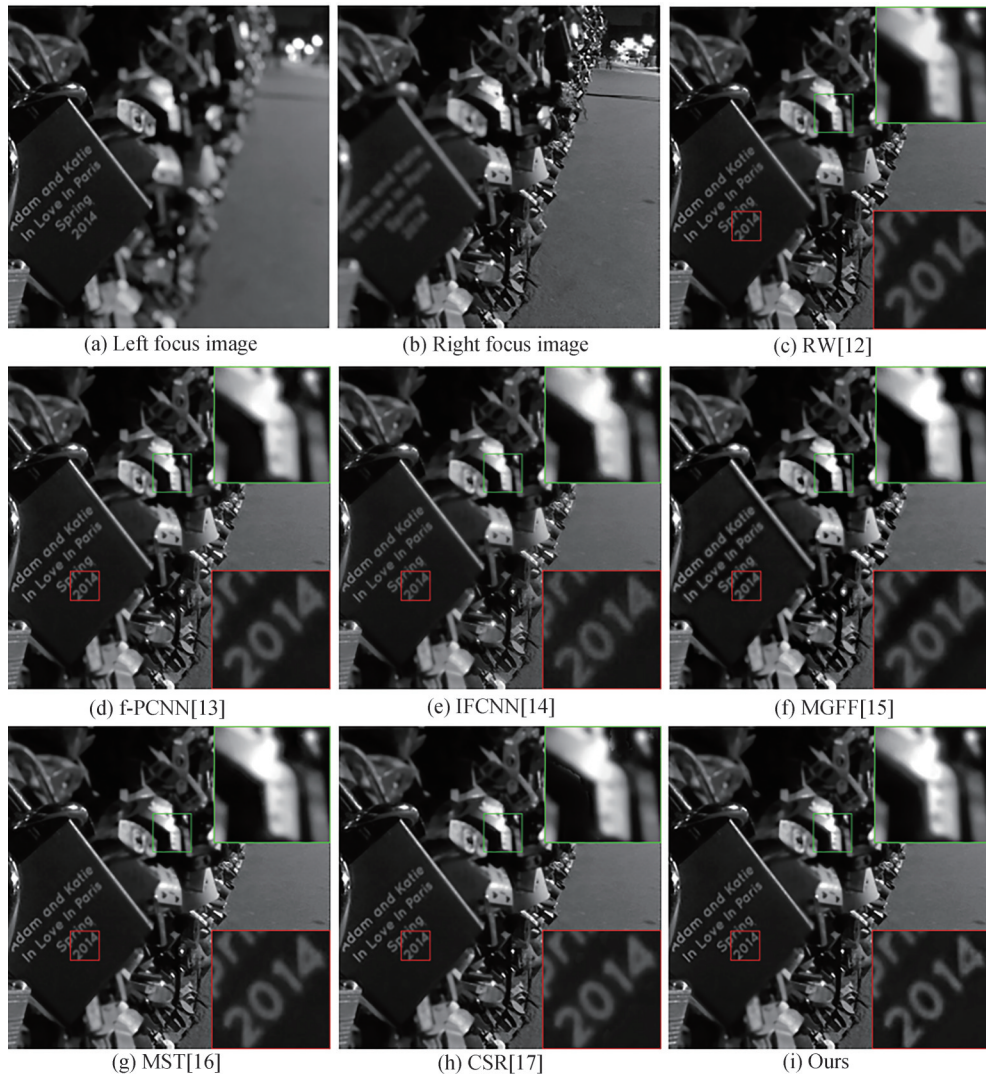


图 10 LoveLocks 图像融合结果对比
Fig. 10 LoveLocks image fusion results comparison

表 1 LoveLocks 图像在 7 种融合算法上的客观指标对比

Table 1 Objective index comparison of LoveLocks image on seven fusion algorithms

	SF \uparrow	Var \uparrow	Q_{cv} \downarrow	Q_{cb} \uparrow	Q_{abf} \uparrow	MI \uparrow	EN \uparrow	EI \uparrow	CE \downarrow	AG \uparrow
RW	9.09	46.31	92.78	0.73	0.775	5.27	7.089	32.02	0.098	29.28
f-PCNN	9.49	48.90	212.27	0.68	0.756	4.38	7.180	34.11	0.011	31.22
IFCNN	9.09	48.47	137.86	0.59	0.602	4.05	7.175	32.75	0.014	29.95
MGFF	8.77	51.02	412.63	0.59	0.705	3.71	7.148	33.11	0.092	30.34
MST	9.54	49.04	111.51	0.70	0.762	4.63	7.164	34.05	0.014	31.15
CSR	9.08	48.34	90.76	0.69	0.733	4.87	7.145	32.15	0.010	29.46
M&P	9.51	50.28	92.73	0.77	0.763	5.65	7.151	33.19	0.008	30.38

Note: Red represents the optimal value, blue represents the suboptimal value, "M&P" represents the proposed algorithm

域,可以看出RW方法和CSR方法的融合结果具有明显的伪影,推测正是因为这些伪影导致 Q_{cv} 值的异常。

为了使7种方法融合结果指标对比更为公正客观,对10组融合结果的评价指标求平均值,如表3所示。由表3可知,本文算法在4类客观指标中均有较好表现,在空间频率、 Q_{cb} 、互信息量、信息熵、边缘强度和交叉熵这6个指标中取得了最优值,其中空间频率和互信息量表现尤为突出,说明经本算法生成的全聚焦图像具有更清晰的细节和纹理且对重聚焦图像具有更强的信息保留能力。本文算法在剩下4个指标中均为次优

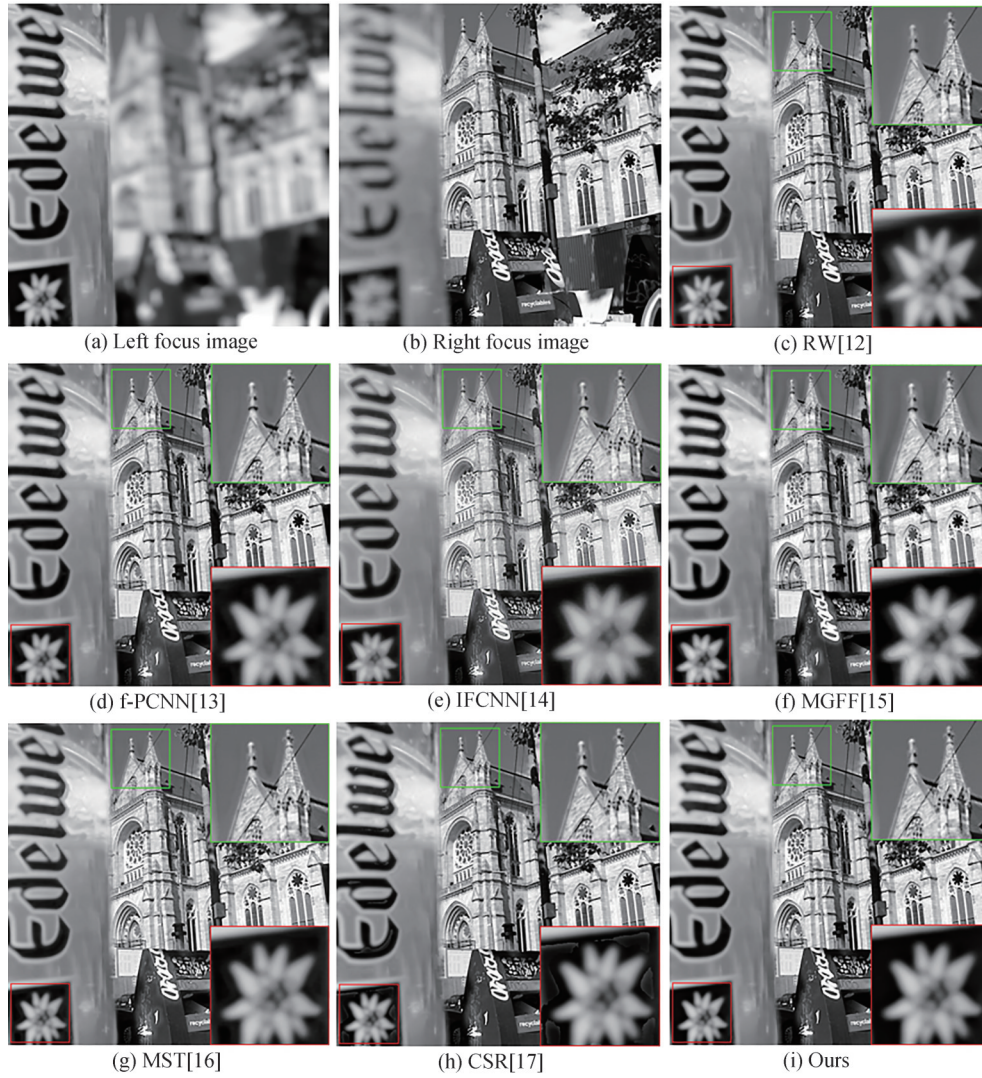


图 11 Edelweiss 图像融合结果对比

Fig. 11 Edelweiss image fusion results comparison

表 2 Edelweiss 图像在 7 种融合算法上的客观指标对比

Table 2 Objective index comparison of Edelweiss images on seven fusion algorithms

	SF \uparrow	Var \uparrow	Q_{cv} \downarrow	Q_{cb} \uparrow	Q_{abf} \uparrow	MI \uparrow	EN \uparrow	EI \uparrow	CE \downarrow	AG \uparrow
RW	21.80	71.13	84.85	0.80	0.838	4.89	7.955	77.66	0.088	69.26
f-PCNN	22.02	70.65	118.67	0.75	0.821	3.37	7.963	79.11	0.070	70.54
IFCNN	22.01	68.36	210.74	0.56	0.693	2.96	7.948	77.96	0.057	69.57
MGFF	20.03	72.81	154.58	0.64	0.757	2.75	7.930	71.89	0.105	64.20
MST	21.95	70.35	96.31	0.72	0.823	3.13	7.918	78.13	0.100	69.71
CSR	21.78	71.99	87.15	0.81	0.817	4.42	7.963	78.12	0.074	69.55
M&P	22.09	72.69	88.85	0.87	0.839	5.51	7.964	79.10	0.076	70.46

Note: Red represents the optimal value, blue represents the suboptimal value, "M&P" represents the proposed algorithm

值,其中,MGFF方法虽然具有最高的方差,但在3个视觉指标中的表现较差,这可能是由于MGFF方法所产生的融合图像会在边界处会产生伪影,从而导致其方差虚高;两个视觉指标 Q_{cv} 和 Q_{abf} 与最优指标非常贴近且远高于排名第三的指标。综上所述,经本文算法生成的多聚焦图像与其他6种方法相比具有清晰度高、视觉效果优秀、图像对比度高等特点。

表4给出了7种方法在10组多聚焦融合实验中所需要的平均时间成本。可以看出,MST、MGFF以及

表3 7种融合算法的平均客观指标对比
Table 3 Average objective index comparison of seven fusion algorithms

	SF ↑	Var ↑	Q_{cv} ↓	Q_{cb} ↑	Q_{abf} ↑	MI ↑	EN ↑	EI ↑	CE ↓	AG ↑
RW	10.47	60.34	38.60	0.776	0.765	6.14	7.551	37.23	0.036 8	33.98
f-PCNN	10.61	60.49	81.21	0.724	0.735	5.05	7.552	37.91	0.026 4	34.61
IFCNN	10.38	60.27	72.74	0.594	0.591	4.65	7.545	36.52	0.021 8	33.36
MGFF	9.86	62.36	169.86	0.634	0.681	4.34	7.543	35.43	0.062 3	32.39
MST	10.68	60.72	48.59	0.740	0.751	5.20	7.542	37.89	0.031 9	34.59
CSR	10.35	60.65	34.42	0.722	0.714	5.53	7.551	36.28	0.021 5	33.12
M&P	10.75	60.79	35.92	0.795	0.761	6.27	7.556	37.93	0.021 4	34.59

Note: Red represents the optimal value, blue represents the suboptimal value, "M&P" represents the proposed algorithm

表4 7种融合算法的平均时间成本
Table 4 Average time cost of seven fusion algorithms

Algorithm	RW	f-PCNN	IFCNN	MGFF	MST	CSR	M&P
Time cost	22.51	532.14	11.85	3.53	0.05	228.81	204.18

Note: Red represents the optimal value, blue represents the suboptimal value, "M&P" represents the proposed algorithm

IFCNN 三种方法均具有较低的时间成本,但在客观指标上的表现却不佳:MST方法具有最低的时间成本,但仅取得了空间频率次优的成绩,MGFF方法仅在方差上获得了最优值,IFCNN方法在10项客观指标中均表现平庸。本文方法的时间成本低于CSR方法和f-PCNN方法,综合看来,本文方法取得了良好的融合效果,但是由于引入核主成分分析对各层特征系数进行融合从而导致了时间成本较高。

4 结论

本文提出了一种基于多尺度潜在低秩分解的光场全聚焦图像融合方法。通过对4D光场图像进行数字重聚焦,将焦点聚集到图像中的各个目标上从而得到重聚焦图像,然后将重聚焦图像作为输入图像来实行全聚焦融合。本文引入多尺度潜在低秩分解对重聚焦图像进行分解,使用局部双区域梯度差值加权平均算法和可迭代引导滤波视觉显著性提取算法来计算各层特征信息,最后使用基于核主成分分析的特征系数融合算法生成聚焦决策图并引导全聚焦图像的生成。实验结果表明,经本文方法生成的光场全聚焦图像具有良好的视觉效果以及更高的空间分辨率;与近年提出的基于传统或深度学习的多聚焦图像融合算法相比,本文算法在所选4类客观评价指标中均有良好表现,验证了其有效性。实验过程中选用灰度图像作为输入图像,目的是为了减少计算量和时间成本,本文算法也同样适用于彩色光场图像的全聚焦融合。此外,该算法也存在不足之处,在各重聚焦图像聚焦边界区域的处理上,仅采用引导滤波对其进行平滑融合,并未做出进一步的优化,今后将对模糊边界区域的融合规则进行研究,使生成的全聚焦图像具有更好的视觉表现。

参考文献

- [1] NG R, LEVOY M, BREDIF M, et al. Light field photography with a hand-held plenoptic camera[J]. Technical Report CTSR, 2005:2.
- [2] PARK J H, LEE S K, JO N Y, et al. Light ray field capture using focal plane sweeping and its optical reconstruction using 3D displays[J]. Optics Express, 2014, 22(21): 25444-25454.
- [3] YU L A, LEI W A, JC A, et al. Multi-focus image fusion: a survey of the state of the art[J]. Information Fusion, 2020, 64: 71-91.
- [4] WU Y, WANG Y, LIANG J, et al. Light field all-in-focus image fusion based on spatially-guided angular information[J]. Journal of Visual Communication and Image Representation, 2020, 72: 102878.
- [5] XIE Yinxian, WU Yingchun, WANG Yumei, et al. Light field all-in-focus image fusion based on wavelet domain definition evaluation[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1848-1854.
谢颖贤, 武迎春, 王玉梅, 等. 基于小波域清晰度评价的光场全聚焦图像融合[J]. 北京航空航天大学学报, 2019, 45(9): 1848-1854.
- [6] SU Boni. Light field all-in-focus image generation method based on gradient domain[J]. Journal of Southwest University (Natural Science), 2020, 42(10): 174-182.
苏博妮. 基于梯度域的光场全聚焦图像生成方法[J]. 西南大学学报(自然科学版), 2020, 42(10): 174-182.

- [7] BAI X, LIU M, CHEN Z, et al. Multi-focus image fusion through gradient-based decision map construction and mathematical morphology[J]. *IEEE Access*, 2016, 4: 4749-4760.
- [8] MA J, ZHOU Z, WANG B, et al. Infrared and visible image fusion based on visual saliency map and weighted least square optimization[J]. *Infrared Physics & Technology*, 2017, 82: 8-17.
- [9] SCHÖLKOPF B, SMOLA A, MÜLLER K R. Nonlinear component analysis as a kernel eigenvalue problem[J]. *Neural Computation*, 1998, 10(5): 1299-1319.
- [10] CAO L J, CHUA K S, CHONG W K, et al. A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine[J]. *Neurocomputing*, 2003, 55(1-2): 321-336.
- [11] HONAUER K, JOHANNSEN O, KONDERMANN D, et al. A dataset and evaluation methodology for depth estimation on 4D light fields[C]. *Proceedings of Asian Conference on Computer Vision*. Springer, Cham, 2016: 19-34.
- [12] MA J, ZHOU Z, BO W, et al. Multi-focus image fusion based on multi-scale focus measures and generalized random walk[C]. *IEEE Proceedings of Chinese Control Conference (CCC)*, 2017: 5464-5468.
- [13] PANIGRAHY C, SEAL A, MAHATO N K. Fractal dimension based parameter adaptive dual channel PCNN for multi-focus image fusion[J]. *Optics and Lasers in Engineering*, 2020, 133(6):106141.
- [14] YU Z A, YU L B, PENG S C, et al. IFCNN: A general image fusion framework based on convolutional neural network [J]. *Information Fusion*, 2020, 54:99-118.
- [15] BAVIRISETTI D P, XIAO G, ZHAO J, et al. Multi-scale guided image and video fusion: a fast and efficient approach [J]. *Circuits Systems and Signal Processing*, 2019, 38(12): 5576-5605.
- [16] YU L, LIU S, WANG Z. A general framework for image fusion based on multi-scale transform and sparse representation [J]. *Information Fusion*, 2015, 24: 147-164.
- [17] LIU Y, CHEN X, WARD R K, et al. Image fusion with convolutional sparse representation[J]. *IEEE Signal Processing Letters*, 2016, 23(12): 1882-1886.
- [18] MOUSNIER A, VURAL E, GUILLEMOT C. Partial light field tomographic reconstruction from a fixed-camera focal stack[J]. *arXiv preprint arXiv:1503.01903*, 2015.
- [19] MA J, MA Y, LI C. Infrared and visible image fusion methods and applications: a survey[J]. *Information Fusion*, 2019: 153-178.
- [20] CVEJIC N, OZA A, BULL D, et al. A similarity metric for assessment of image fusion algorithms [J]. *International Journal of Signal Processing*, 2005, 2(3): 178-182.
- [21] YIN C, BLUM R S. A new automated quality assessment algorithm for image fusion[J]. *Image & Vision Computing*, 2009, 27(10): 1421-1432.
- [22] HAO C, VARSHNEY P K. A human perception inspired quality metric for image fusion based on regional information [J]. *Information Fusion*, 2007, 8(2): 193-207.

Light Field All-in-focus Image Fusion Based on MDLatLRR and KPCA

HUANG Zefeng, YANG Shen, DENG Huiping, LI Qingson

(*School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China*)

Abstract: The imaging of a light field camera can retain the spatial and angular information of light, therefore, different from traditional two-dimensional imaging, the light field camera can capture the light field directly in one shot, but it will sacrifice the spatial resolution and angular resolution of the image, so the quality of the image obtained is lower than that of the image generated by the native image sensor. This problem has prevented the application of light field imaging from gaining popularity. The development of multi-focus image fusion technology and digital refocusing technology provides ideas for improving the resolution of light field imaging. To improve the spatial resolution of light field imaging, we propose a full-focus fusion algorithm of light field image based on multi-scale latent low-rank decomposition and kernel principal component analysis by combining digital refocusing of light field image with multi-focus image fusion. First, by reprojecting the light field, the light is projected from the original focusing plane to the refocusing plane to generate a refocusing image with the focusing region, defocus region and blurred boundary region. After digitally refocusing, the spatial resolution of the light field image located in the focusing area is greatly improved. To extract the focus area accurately, the multi-level image

decomposition method based on latent low-rank representation is used to decompose each refocusing image into a base layer and several saliency layers. Then, a two-region image sharpness extraction algorithm is used to calculate the image sharpness of the base layer, and a multi-scale saliency extraction algorithm is used to extract the visual saliency of the saliency layer's gradient domain. Secondly, the feature coefficient matrices of the base layer and each saliency layer are reshaped and concatenated. The kernel principal component analysis is used for dimension reduction fusion to obtain the fusion feature coefficient matrix which retains the feature information of both the base layer and the salient layer. Finally, the initial focusing decision map was generated by comparing the fusion feature coefficient matrix corresponding to each refocusing image. The small structure removal is applied to the initial decision map to eliminate the influence caused by image noise, and guided filtering is used to process the initial focusing decision map to generate the final focusing decision map, which solves the fusion problem at the fuzzy boundary, so that the fused image is smoother at the boundary of the focusing decision map and has a better visual effect. In order to verify the effectiveness of the proposed method, two sets of experiments are carried out: the full-focus fusion experiment and the multi-focus image fusion experiment. In the full-focus fusion experiment, several kinds of light field data, including objects located at different depths and flat backgrounds were selected from the HCI dataset, which can test the full-focus fusion ability of the proposed algorithm for complex textures and multi-objects. In the multi-focus image fusion experiment, six traditional multi-focus image fusion methods are selected for comparison with the proposed method, the experimental data are from the LFSD dataset, and ten objective evaluation indicators are selected for comparison. The experimental results show that the algorithm has better performance in visual effect and edge information richness compared with traditional methods. Subjectively, the full-focus images fused by the proposed method have higher spatial resolution and detailed texture compared with the original light field imaging. The resulting full-focus image is more in line with the visual perception of human eyes, objectively. Compared with the other six multi-focus image fusion methods, the multi-focus image generated by the proposed algorithm has the characteristics of high definition, excellent visual effect and high image contrast.

Key words: Light field; All-in-focus image fusion; Digital refocusing; MDLatLRR; Kernal principal component analysis

OCIS Codes: 100.2000; 100.2980; 110.2960; 110.3080