

## 融合一致性与差异性约束的光场深度估计

何泽阳<sup>1,2</sup>, 邓慧萍<sup>1,2</sup>, 向 森<sup>1,2</sup>, 吴 谨<sup>1,2</sup>

- (1. 武汉科技大学 信息科学与工程学院, 湖北 武汉 430081;
2. 武汉科技大学 冶金自动化与检测技术教育部工程研究中心, 湖北 武汉 430081)

**摘 要:** 光场图像深度估计是光场三维重建、目标检测、跟踪等应用中十分关键的技术。虽然光场图像的重聚焦特性为深度估计提供了非常有用的信息,但是在处理遮挡区域、边缘区域、噪声干扰等情况时,光场图像深度估计仍然存在很大的挑战。因此,提出了一种基于极平面图(Epipolar plane image, EPI)斜线像素一致性和极平面图区域差异性的深度估计算法用于解决遮挡和噪声问题。EPI斜线像素的一致性采用旋转线性算子(Spinning linear operator, SLO)的颜色熵度量,能够提高深度图边缘的准确性以及抗噪能力;EPI区域的差异性采用旋转平行四边形算子(Spinning parallelogram operator, SPO)的卡方 $\chi^2$ 度量,能够提高深度图深度渐变区域的准确性,并使用置信度加权的方法将两种度量进行融合,可以减少遮挡区域和噪声的干扰。另外,充分利用像素邻域的颜色相似性,使用引导保边滤波器和马尔科夫随机场(Markov random field, MRF)全局优化策略进行后处理,进一步减少深度图的边缘错误,得到遮挡边缘准确的深度图。在HCI光场数据集上进行了实验,并与经典光场深度估计算法进行了对比,结果表明该算法在主观质量和客观指标两方面都有明显提升。

**关键词:** 光场图像; 深度估计; EPI; 像素一致性

中图分类号: TP391 文献标志码: A DOI: 10.3788/IRLA20210021

## Light field depth estimation of fusing consistency and difference constraints

He Zeyang<sup>1,2</sup>, Deng Huiping<sup>1,2</sup>, Xiang Sen<sup>1,2</sup>, Wu Jin<sup>1,2</sup>

- (1. School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China;
2. Engineering Research Center for Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China)

**Abstract:** Light field image depth estimation is a critical technique in applications such as light field 3D reconstruction, target detection and tracking. The refocusing property of light field image provided rich information for depth estimation, however it was still challenging in the case of occlusion region, edge region, noise interference, etc. Therefore, a depth estimation algorithm based on the consistency of epipolar plane image (EPI) slant pixels and the difference of epipolar plane image regions was proposed to solve the occlusion and noise problems. The consistency of EPI slant pixels was adopted by the spinning linear operator (SLO) color entropy metric, which could improve the accuracy of depth map edges as well as the noise immunity; the

收稿日期:2021-05-25; 修订日期:2021-06-20

基金项目:国家自然科学基金(61702384, 61502357); 湖北省自然科学基金(2015CFB365)

作者简介:何泽阳,男,硕士生,主要从事图形图像处理方面的研究。

导师简介:邓慧萍,女,副教授,主要从事3D视频处理方面的研究。

difference of EPI regions was measured by the chi-square  $\chi^2$  metric of the spinning parallelogram operator (SPO), which could improve the accuracy of the depth gradient area of the depth map, and the two metrics were fused with confidence weighting, which could reduce the interference of occluded regions and noise. In addition, the color similarity of the pixel neighborhood was fully utilized, and post-processing was performed with a guided edge-preserving filter and Markov random field (MRF) global optimization strategy to further reduce the edge error of the depth map to obtain an accurate depth map with occluded edges. Experiments were conducted on the HCI light field data set and compared with the classic light field depth estimation algorithm. The results show that the algorithm has a significant improvement in both subjective quality and objective indicators.; that the proposed method outperformed state-of-the-art depth estimation methods. Compared with the classical light field depth estimation algorithm, proposed method had significant improvement on both subjective and objective qualities.

**Key words:** light field images; depth estimation; epipolar plane image(EPI); pixel consistency

## 0 引言

4D 光场 (LF) 图像<sup>[1]</sup> 是一种同时记录场景中的强度信息和光线方向的图像。早些时期, 人们通过密集相机阵列获取的 4D 光场图像, 但这种方法成本较高耗时较长, 图像基线较大。近年来, 随着商用光场相机 Lytro<sup>[2]</sup> 和 Raytrix<sup>[3]</sup> 的出现, 可以更快速便捷的得到基线较小的 4D 光场图像。4D 光场图像与传统 2D 图像相比有着更加广泛的应用, 光场深度估计作为光场图像处理中的关键应用, 可以使用数字重聚焦<sup>[4]</sup> 的结果进行处理, 是图像超分辨率、三维重建、3D 点云<sup>[5]</sup> 中的关键技术。

深度估计是指通过图像中的信息估计对应场景的深度信息, 传统图像的深度估计通常通过立体匹配的思想计算深度, 而光场图像拥有不同视角的光线方向信息, 可以通过光场图像的不同特性计算场景深度。目前, 光场深度估计方法一般分为以下几类: 基于子孔径图像的多视角立体匹配算法、基于重聚焦图像和角度块的算法、基于极平面图 (Epipolar plane image, EPI)<sup>[6]</sup> 的算法。

EPI 作为光场图像特有的一种可视化方法, 同时包含空间域与角度域的信息, 有利于解决深度估计的遮挡问题。文中以 EPI 斜线像素一致性作为 EPI 视差图的判断准则, 建立了一种基于旋转线性算子 (Spinning linear operator, SLO) 的代价函数, 并利用 RGB 颜色熵的代价函数对斜线像素的一致性进行衡量。同时, 充分利用 SPO 的优点, 采用基于代价成本置信度的融合方法, 将 SLO 的代价成本与 SPO 代价成本进行加权融合。最后, 利用引导保边缘滤波器和

全局优化方法, 进一步提高深度图准确性。提出的算法与传统光场深度估计算法在 HCI 数据集上进行了实验验证和比较, 结果表明在边缘处理和噪声鲁棒性方面, 文中算法表现出更好的性能。

文中的主要贡献包括:

(1) 提出了一种基于颜色熵和旋转线性算子的 EPI 斜线像素一致性代价函数用于深度估计。该方法可以准确的计算遮挡边缘, 并对噪声鲁棒性较强;

(2) 引入了一种置信度加权的融合方法, 将 EPI 斜线像素一致性代价函数和基于 SPO 的 EPI 区域差异进行代价函数结合, 形成代价互补。之后使用引导滤波和基于马尔科夫随机场 (Markov random field, MRF) 的全局优化策略进行后处理, 得到的深度图在 HCI 数据集上的综合表现优于目前经典算法。

## 1 光场深度估计研究现状

传统的光场深度估计方法根据光场可视化方法的不同, 主要包括基于多视角立体匹配的算法、基于重聚焦图像和角度块的算法、基于 EPI 的算法。

基于多视角立体匹配的算法基本思想来自于传统 2D 立体匹配, 通过光场子孔径图像中相邻视角图像的视差关系, 进行深度估计。Jeon<sup>[7]</sup> 提出了一种基于相移的亚像素多视角匹配算法用于光场深度估计。Chen<sup>[8]</sup> 设计了一种基于双边统计的光场立体匹配算法。Zhang<sup>[9]</sup> 提出了一种基于微透镜匹配的光场信息恢复框架, 可用于深度估计。Fan<sup>[10]</sup> 提出了一种基于光场结构特性优化的多视点匹配光场深度估计算法。然而基于多视角立体匹配的算法与传统的立

体匹配算法类似,对遮挡区域的处理效果不好,很难处理多遮挡的场景,对噪声也比较敏感。

基于重聚焦和角度块的深度估计算法的主要思想是利用光场图像可以先拍照再聚焦的特性,获得全聚焦图像和焦堆栈,然后匹配焦堆栈角块和子孔径图像角块寻求正确的聚焦度,最后根据聚焦度与深度的对应关系求得准确的深度。这类方法利用了光场图像的重聚焦特性,能有效的处理遮挡和噪声问题。Tao<sup>[11-12]</sup>提出了一种以融合离焦线索和匹配线索作为代价函数的算法,并在之后加以阴影约束优化代价,以获得正确的深度图。Wang<sup>[13]</sup>提出了一种基于角度块与空间块相似性的算法,对遮挡的鲁棒性十分出色,但初始结果处理较差,依赖于后处理和优化。Willem<sup>[14-15]</sup>提出了一种基于角度熵的深度估计算法,并采用了全局优化的模型,得到了很好的效果。Guo<sup>[16]</sup>设计了一种基于多方向角相干的深度估计算法,根据角相干方向与空间中遮挡方向相同的原理处理遮挡问题。Sheng<sup>[17]</sup>设计了一种基于角度块特性的积分引导滤波器,可以应用到其他光场深度估计框架中处理遮挡问题。Chen<sup>[18]</sup>提出了一种基于超像素正则化的深度估计算法,在处理部分遮挡场景时展现出杰出的效果。然而这类算法比较依赖于焦堆栈角块中的图像信息,当焦堆栈角块信息不足,例如处理完全遮挡场景和深度变化不明显区域时,表现较差。

EPI 作为光场图像的一种特殊的可视化表示方法,最早由 Bolles<sup>[6]</sup>提出。基于 EPI 的光场深度估计

将光场图像的表达形式从普通的子孔径图像转换为 EPI,从而在一张图上同时获取空间坐标和角度坐标,将视差具体为 EPI 斜线斜率。图 1 为 EPI 的获取过程两种表示方法,上半为平面表示,下半为 3D 立体表示。由于这类方法不是直接从图像空间关系入手,会减少遮挡物体的影响。Wanner<sup>[19-20]</sup>提出的结构张量代价在早期研究中被广泛应用,其泛用性和基本性能很强,但也十分依赖于高分辨率光场图像,对遮挡和噪声处理的效果也不够优秀。Kim<sup>[21]</sup>提出的从细到粗的深度估计框架,以边缘点为支撑点做对应 EPI 线的一致性检测,以其深度值做深度扩散得到最终结果。Huang<sup>[22]</sup>在 Kim 的基础上更改了代价,提出了三种线索结合的代价度量斜线一致性,提高了边缘效果。Zhang<sup>[23]</sup>设计了一种旋转平行四边形算子用于 EPI 斜线左右区域的差异比较,以此差异为依据计算深度图。Sheng<sup>[24]</sup>在 Zhang 的基础上提出了基于多方向 EPI 的算法,进一步处理了遮挡问题。Li<sup>[25]</sup>提出了一种基于重聚焦和 EPI 相结合的深度估计算法,通过正确聚焦的 EPI 线垂直的原理计算深度,对遮挡处理较好。Schilling<sup>[26]</sup>设计了一种基于 EPI 线遮挡的优化算法,对 SPO 算法在某些边缘处处理不好的缺点进行了优化,达到了优秀的效果。其中,SPO 算法是目前基于 EPI 的光场深度估计算法中性能最好的一种算法,其思想是从 EPI 斜线斜率与深度对应关系出发,通过 EPI 斜线两侧区域差异性作为 EPI 斜线选择标准,利用直方图卡方距离度量差异性,得到深度

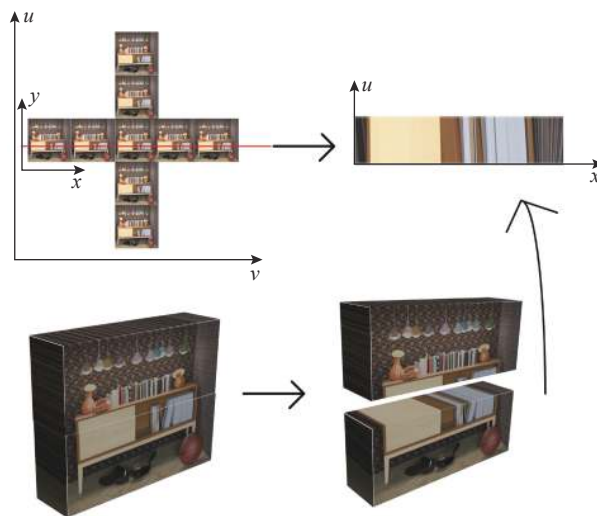


图 1 光场及 EPI 获取示意图(上),将光场图片叠加后剪切示意图(下)

Fig.1 Schematic diagram of light field and EPI acquisition (top), and diagram after overlaying light field images (bottom)

图。该算法对大多遮挡和噪声处理效果较好,但基本忽略了 EPI 斜线上的像素一致性,所以对某些遮挡边缘区域会产生错误估计。

## 2 基于 SLO 和 SPO 的光场深度估计

### 2.1 基于 EPI 的深度图估计

4D 光场图像可表示为:  $L(x,y,u,v)$ , 其中  $x, y$  为光场的空间坐标,  $u, v$  为光场的角度坐标。EPI 图像是一种包含空间域和角度域信息的光场表示形式。因此,当固定光场图像的  $y$  和  $v$ , 改变  $x$  和  $u$  时,就可以得到以  $x$  和  $u$  为长和宽的横向 EPI, 表示为:

$$U_{y,v}(x,u) = L(x,y^*,u,v^*) \quad (1)$$

同样,当假定光场图像的  $x$  和  $u$  固定, 只改变  $y$  和  $v$  可以产生纵向的 EPI, 表示为:

$$U_{x,u}(y,v) = L(x^*,y,u^*,v) \quad (2)$$

视差是真实世界中的一点在不同视角的子孔径图像中位置的差异。在 EPI 中, 视差可以通过同一个像素点组成的 EPI 斜线倾斜角度来表示。如图 2 所示, 以点  $P$  为中心包含一组旋转的 EPI 斜线, 当倾斜

角为  $\theta$  的斜线  $L_1$  为正确的 EPI 斜线 (所对应的斜率为该点的深度), 其深度  $z$  就可以计算为:

$$z = \frac{f \cdot \Delta u}{\Delta x} = -\frac{f}{\tan \theta} \quad (3)$$

式中:  $f$  为相机焦距;  $\Delta u$  为相邻视角图像的度量距离;  $\Delta x$  为同一个点在两个相邻不同视角中的视差值;  $\tan \theta$  表示 EPI 中相同像素组成的斜线斜率。通过公式 (3) 求解真实深度的问题就被转变成了求解角  $\theta$  值的问题。

如图 2 中  $L_1$  所示, 组成这条斜线的每个点的像素值基本一致, 所以其斜线一致性相比于错误的 EPI 斜线  $L_2, L_3$  来说更大。另一方面, 由  $L_1$  分割的两侧区域  $\lambda_1$  和  $\lambda_2$  的像素分布可以看出, 相比于错误 EPI 斜线  $L_2, L_3$  两侧区域, 其区域差异性更大。

因此, 文中从 EPI 斜线斜率的角度出发, 通过 EPI 斜线像素一致性和 EPI 斜线两侧区域差异性两个方面的约束来确定正确的 EPI 斜线, 提出基于 SLO 的 EPI 斜线一致性代价函数和基于 SPO 的 EPI 区域差异性代价函数来估计光场的深度图。如图 3 所示

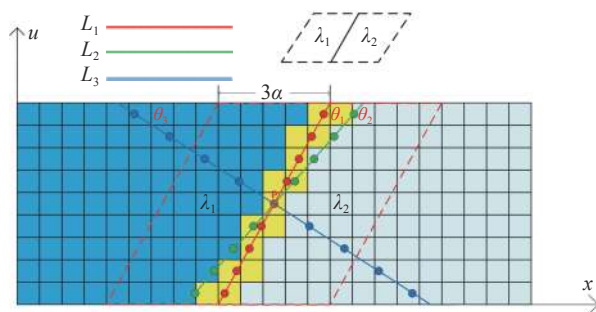


图 2 SPO、SLO 示意图

Fig.2 Schematic diagram of SPO and SLO

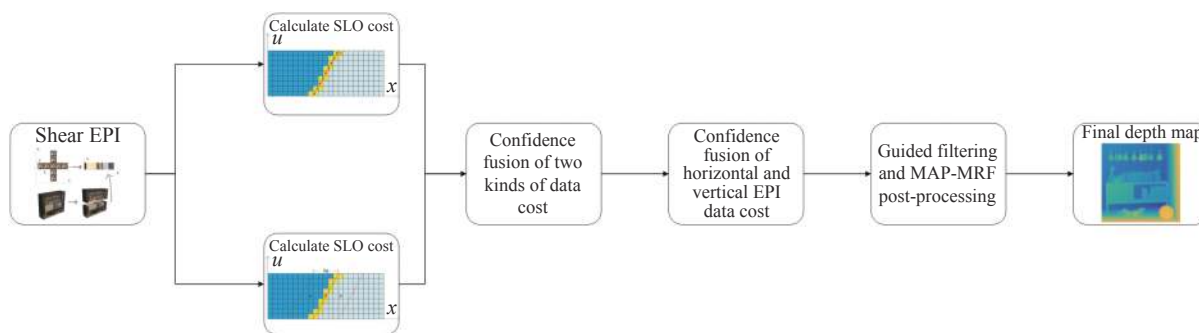


图 3 光场深度估计流程图

Fig.3 The flow chart of light field depth estimation

为整体流程: 首先对子孔径图像进行剪切得到 EPI, 然后分别通过基于 SLO 的 EPI 斜线一致性代价函数和基于 SPO 的 EPI 区域差异性代价函数得到两种代价成本, 再通过自适应置信度权重融合两种代价, 并置信度融合纵横 EPI 代价, 最后采用引导滤波器和 MAP-MRF 进行后处理, 得到最终深度图。

## 2.2 基于 SLO 的 EPI 斜线一致性代价函数

为了得到 EPI 斜线, 文中建立了新型的旋转线性算子:

$$R(x^*, \theta_N) = \{E(x^* + \Delta x_N, u) \mid u = 1, 2, \dots\} \quad (4)$$

$$\Delta x_N = \{(u - u^*) (1 / \tan \theta_N) \mid u = 1, 2, \dots\} \quad (5)$$

式中:  $R(x^*, \theta_N)$  表示对应  $\theta_N$  的假设 EPI 斜线;  $x^*$  为中心视角中的所求像素点的水平坐标;  $E(x^* + \Delta x_N, u)$  表示在不同的角度坐标下, 假设 EPI 线上的像素在图像中的坐标信息;  $\tan \theta_N$  表示假定 EPI 斜线斜率。

特别地, 为了解决半像素问题, 文中采用了线性插值的方法来计算 EPI 斜线上每一个像素的假定像素值:

$$I(x, u') = \frac{1}{w_1} I_1 + \frac{1}{w_2} I_2 \quad (6)$$

式中:  $I(x, u')$  表示  $E(x^* + \Delta x_N, u)$  坐标位置的假定像素值;  $I_1$  和  $I_2$  表示该假定像素左右相邻的整像素值;  $w_1$  和  $w_2$  表示该像素位置到左右两点的水平距离权重, 根据该式, 距离该像素近的整像素占比大, 距离该像素远的整像素占比小, 可以求得假定的 EPI 斜线像素值。

得到斜线像素值之后, 需要对这组像素值进行一致性衡量, 由于信息熵这类代价函数对比于直接通过像素值进行计算的代价函数, 考虑的是 EPI 斜线的像素复杂程度, 当有噪声干扰使某个像素值变化较大时, 不会对整体斜线的代价函数结果产生过大的影响, 对噪声鲁棒性更强。因此文中设计了基于颜色熵的 EPI 斜线像素一致性代价函数:

$$C_{ei}(P, \theta_N) = - \sum_{u' \in u} h(x, u') \log(h(x, u')) \quad (7)$$

式中:  $h(x, u')$  表示  $I(x, u')$  的像素值概率。

由于颜色熵基本舍弃了像素值本身的信息, 只考虑了概率分布, 对某些特殊像素分布会产生错误衡

量。为了考虑像素值信息, 文中通过计算 EPI 斜线上点的像素值与中心视角像素值差值的均值, 设计了一种基于颜色相似性的自适应权重, 并加以高斯核, 从两方面约束了 EPI 斜线上的像素一致性, 增加代价函数的鲁棒性。表示为:

$$C_e(P, \theta_N) = e^{-\left(\frac{C_{ei}(P, \theta_N)}{2\sigma^2}\right)} w(P, \theta_N) \quad (8)$$

$$w(P, \theta_N) = \frac{e^{-\left(\frac{S_W}{2\sigma^2}\right)}}{N} \quad (9)$$

$$S_W = - \left( \sum_{u' \in u} (I(x, u') - I(x, u^*))^2 \right) \quad (10)$$

式中:  $C_e(P, \theta_N)$  表示最终的单通道熵代价函数;  $C_{ei}(P, \theta_N)$  为公式 (7) 计算得到的初始单通道熵代价函数;  $w(P, \theta_N)$  表示假定 EPI 斜线上所有点与目标像素点的颜色相似性权重;  $N$  为像素个数;  $S_W$  表示 EPI 斜线像素与中心视角像素的差异, 具体表示为所有像素与中心视角像素的像素值差值平方和。当假定 EPI 斜线上像素点差异性越小, 其熵值和颜色差异就越小, 初始颜色熵成本值和权重值就越大, 最终代价值就越大, 达到了双线索互补的效果。

最后充分利用颜色空间信息, 通过 RGB 三通道均值求得最终的颜色熵代价函数:

$$D_e(P, \theta_N) = \frac{C_{eR}(P, \theta_N) + C_{eG}(P, \theta_N) + C_{eB}(P, \theta_N)}{3} \quad (11)$$

## 2.3 基于 SPO 的 EPI 区域差异性代价函数

基于 SPO 的 EPI 区域差异性代价函数主要通过计算 EPI 斜线两侧区域直方图卡方距离作为差异度量, 并依据像素点到直方图的距离加以不同的直方图权重, 得到最终的代价函数。

为了计算 SPO 算法的代价成本, 首先设计算子和相应权重:

$$w_o(i, j) = c \cdot d_\theta(i, j) \cdot e^{-\frac{d_\theta^2(i, j)}{2\sigma^2}} \quad (12)$$

式中:  $d_\theta(i, j) = i - (x_r + (j - u_r) \cdot \tan \theta)$  为 EPI 同一行中所选取的深度测量点与其左右邻域中某一点的水平距离,  $(x_r + (j - u_r) \cdot \tan \theta)$  为假定 EPI 斜线与该点交点的水平位置;  $\alpha$  为衡量 SPO 宽度的尺度参数;  $c$  为正则化系数。实验中, 一般取 SPO 的大小为  $u \cdot 3\alpha$ , 即高为角度分辨率  $u$ , 宽为尺度参数  $\alpha$  的三倍。

之后,通过直方图的卡方距离<sup>[27]</sup>衡量假定 EPI 斜线左右领域的颜色差异:

$$\chi^2(a_\theta, b_\theta) = \sum_i \frac{(a_\theta(i) - b_\theta(i))^2}{a_\theta(i) + b_\theta(i)} \quad (13)$$

式中:  $a_\theta(i)$  和  $b_\theta(i)$  为两侧区域的颜色直方图。如果  $\chi^2$  的值很大,即表示两侧区域差异大,所以选取的 EPI 斜线就越正确。按照该原理,可以得到代价成本最大的 EPI 斜线斜率,从而计算正确深度值。特别的,该算法利用卷积的计算方法计算  $\chi^2$  距离,提高了计算效率。最终得到了 SPO 算法的最终直方图卡方距离代价函数:

$$D_s(P, \theta_N) = \chi^2(a_{\theta_N}, b_{\theta_N}) \quad (14)$$

SLO 代价函数和 SPO 代价函数都值分别考虑了对应的 EPI 斜线衡量特性,如图 4 为 HCI 数据集 box 的一张横向 EPI 对的代价成本分布, SPO 求得最大标签为 68, SLO 求得最大标签为 69, GT 标签值为 68.5317。可以看出两种代价函数都能分别计算出正确的 EPI 斜线。

然而,由于两种代价函数都只考虑了一个方面,所以各有缺点。基于 SLO 的斜线一致性代价函数在深度渐变区域表现较差,会检测出错误边缘。而基于 SLO 的区域差异性代价函数在边缘判定和噪声鲁棒性方面表现较差,边缘不够准确噪点过多。如图 5 所示,一张横向 EPI 上 SPO 和 SLO 两种代价函数分别会在某些点计算正确或计算错误。而整体效果如图 6 所示, boxes、dino 图像用 SPO 代价函数与 SLO

代价函数直接 WTA 选择和进行不做后处理的结果比较,从结果可以看出 SPO 代价函数在边缘处有模糊现象,而物体内部噪点也较多;而 SLO 代价函数在深度渐变区域分层较为严重,会检测出错误的边缘情况。从结果可以看出, SPO 算法和 SLO 算法各有利弊。

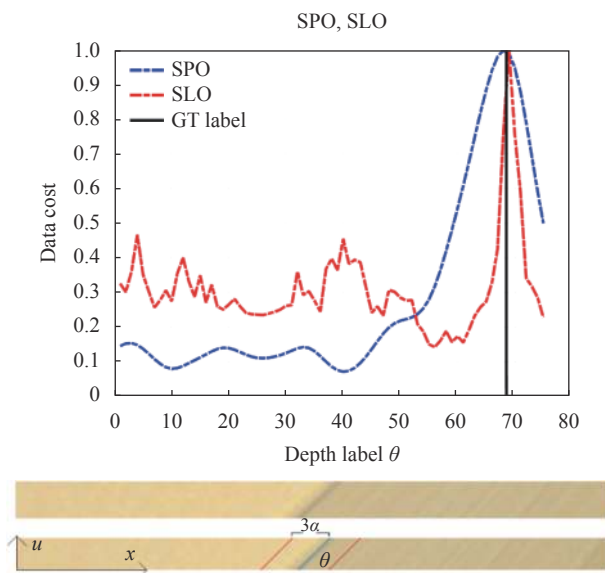
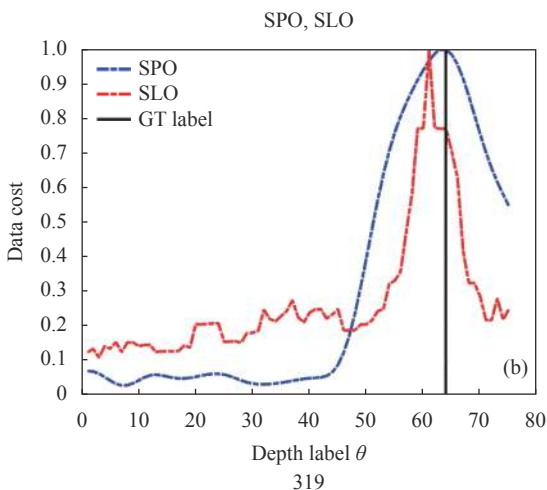
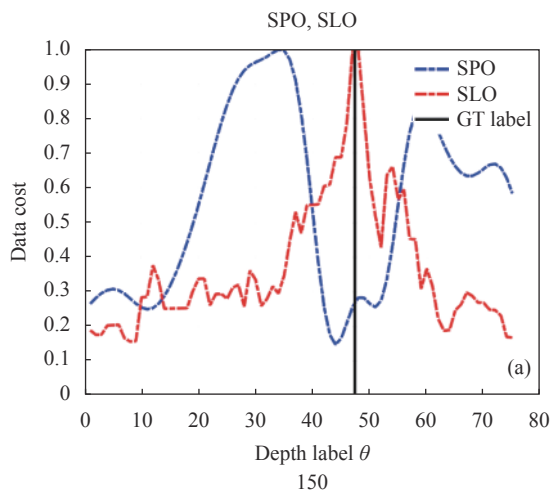


图 4 代价函数准确性分析

Fig.4 Analysis of the accuracy of the cost function

因此,文中提出了一种基于置信度权值的代价融合算法,将两种代价有机的结合在一起,从而使两种算法的形成互补,得到一种新的代价,弥补了两种算法的不足。



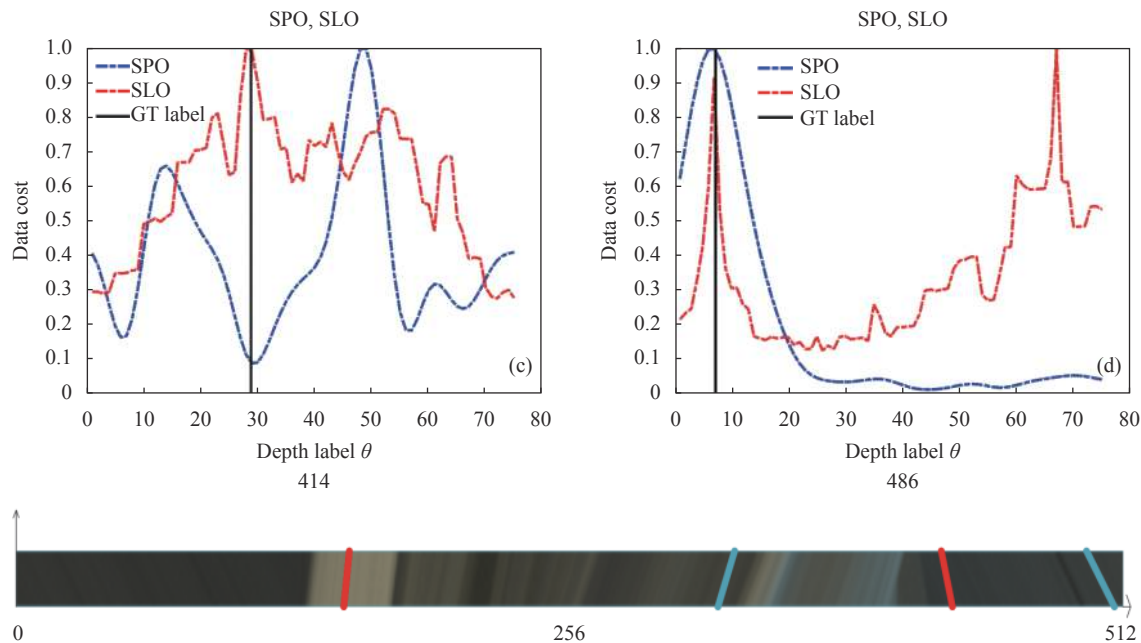


图 5 两种代价函数的准确性对比

Fig.5 Comparison of the accuracy of the two cost functions

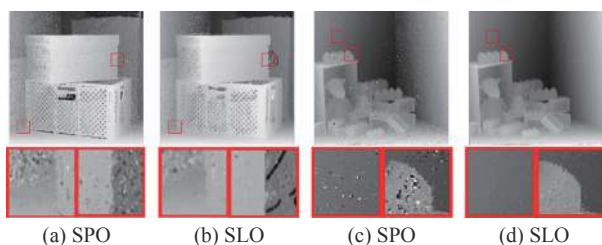


图 6 两种代价函数的深度图实验结果比较

Fig.6 Comparison of experimental results of depth maps with two cost functions

## 2.4 代价融合

首先, 文中设计了一种基于代价成本最大值和平均值比值的置信度权值, 当代价函数计算出的代价成本属于正确 EPI 斜线时, 其分布应呈单峰形状, 所以按照以下公式, 其权值就会更大, 从而更容易被最后的 MRF 函数选择:

$$c = \exp\left(-\frac{\bar{D}/D_{\max}}{2\sigma^2}\right) \quad (15)$$

式中:  $D_{\max} = \max_D D(P, \theta)$ ,  $\bar{D} = \sum_{\theta} D(P, \theta)$ 。

之后通过置信度权值将两种代价结合:

$$D(P, \theta_N) = c_s(P, \theta_N) D_s(P, \theta_N) + c_e(P, \theta_N) D_e(P, \theta_N) \quad (16)$$

特别的, EPI 可以有横向和纵向两种选取方式如

公式 (1)、(2), 所以可以将两个方向的 EPI 代价相结合, 解决不同视角下的遮挡问题:

$$D_{u,v}(x, y, \theta) = c_{y,v^*}(x, u^*) D_{y,v^*}(x, u^*, \theta) + c_{x,u^*}(y, v^*) D_{x,v^*}(y, v^*, \theta) \quad (17)$$

然后对得到的新代价成本采用引导图为中心子孔径的引导滤波器<sup>[28]</sup>进行保边滤波, 得到了初步优化的整体代价成本, 文中公式物理量具体含义如表 1 所示。

表 1 重要符号物理意义

Tab.1 Physical significance of important symbols

Symbol	Physical significance	Symbol	Physical significance
$R(x^*, \theta_N)$	EPI slash	$w_{\theta}(i, j)$	SPO distance weight
$E(x^* + \Delta x_N, u)$	EPI slash pixel coordinate information	$\chi^2(a_{\theta}, b_{\theta})$	SPO histogram $\chi^2$ distance
$\tan \theta_N$	EPI slope	$D_s(P, \theta_N)$	SPO cost function
$I_1, I_2$	Pixel adjacent integer pixel value	$c, c_s, c_e$	Adaptive confidence weight
$w_1, w_2$	Horizontal distance weight	$D(P, \theta_N)$	One-way final cost function
$h(x, u')$	Pixel value probability	$D_{u,v}(x, y, \theta)$	Final cost function
$C_{ei}(P, \theta_N)$	Initial single channel entropy cost	$E_{\text{unary}}$	MRF data cost
$w(P, \theta_N)$	Color similarity weight	$E_s(p, p(\alpha))$	SPO data cost
$C_e(P, \theta_N)$	Final single channel entropy cost	$E_e(p, p(\alpha))$	SLO data cost
$D_e(P, \theta_N)$	SLO color entropy cost function	$E_{\text{binary}}$	MRF smoothing term

### 2.5 基于 MAP-MRF 的全局优化策略

得到初步优化的整体代价成本后,需要对每个深度标签的对应代价成本进行选择,传统方法采用赢家通吃 (Winner take all, WTA) 的算法,直接通过代价值选择出最优结果。文中采用了一种基于马尔科夫随机场框架<sup>[29]</sup>的全局优化模型用于代价优化和选择,定义为:

$$E = \sum_p E_{\text{unary}}(p, \alpha(p)) + \lambda \sum_p \sum_{q \in N(p)} E_{\text{binary}}(p, q, \alpha(p), \alpha(q)) \quad (18)$$

式中:  $p$  为被测量深度的像素点;  $\alpha(p)$  为该像素点的深度标签;  $E_{\text{unary}}$  为马尔科夫最小能量函数的成本项,用于衡量深度标签给予该像素的能量大小;  $N(p)$  为像素  $p$  的邻域像素;  $E_{\text{binary}}$  为马尔科夫最小能量函数的平滑项,通过计算像素  $p$  与其邻域像素的强度一致性 (forces consistency) 得到。

文中提出了一种全局优化模型用于优化每个深度标签代价成本,可展开为:

$$E_{\text{unary}} = c_s E_s(p, p(\alpha)) + c_e E_e(p, p(\alpha)) \quad (19)$$

式中:  $E_s(p, p(\alpha))$  为基于选择平行四边形算子算法 (SPO) 计算得到的最终代价值;  $E_e(p, p(\alpha))$  为基于旋转线性算子所计算得到的最终代价值;  $c_s$  和  $c_e$  分别为两项数据项的置信度权值,用于增加算法的鲁棒性;  $c_s$  和  $c_e$  的具体数值为自适应权值,由公式 (15) 计算得到。

平滑项体现了图像局部空间内的平滑特性,可展开为:

$$E_{\text{binary}}(p, q, \alpha(p), \alpha(q)) = \nabla I(p, q) \min(|\alpha(p) - \alpha(q)|, \tau) \quad (20)$$

式中:  $\nabla I(p, q)$  为邻域像素  $p$  与中心像素  $q$  的强度差异;  $\tau$  为截断值。将初始优化的整体代价通过 MRF 全局模型处理后,既能对代价进行后处理,也能通过最小能量函数选择出正确的深度标签。最终将深度标签对应的深度值映射到图中,就能够得到了文中算法的最终深度图结果。

### 3 实验比较

文中将所提出的算法在 HCI<sup>[30]</sup> 的公开 4D 光场数据集进行了实验验证,并与经典的光场深度估计算法 CAE<sup>[15]</sup>、IGF<sup>[17]</sup>、LF<sup>[7]</sup>、LF\_DC<sup>[12]</sup>、LF\_OOC<sup>[13]</sup>、LF\_PAC<sup>[16]</sup>、MBM<sup>[9]</sup>、POBR<sup>[18]</sup>、SPO<sup>[23]</sup> 进行了比较。

主观质量方面,采取了深度误点图和深度细节图进行对比验证,实验结果如图 7 所示是图片 boxes, cotton, sideboard, dino 测试的误点图结果。从图中可以看出, CAE 算法虽然边缘计算准确,但在深度渐变及背景区域会检测出错误的边缘, LF\_DC 算法处理新 HCI 数据集会产生模糊效应, LF\_OOC 算法在物体内部噪声较多无法滤除, SPO 算法处理某些物体边缘较差,物体内部也会有少量噪点。文中算法误点最少,边缘和背景区域计算也基本正确,对比于其他传统算法基本达到了最好的效果。

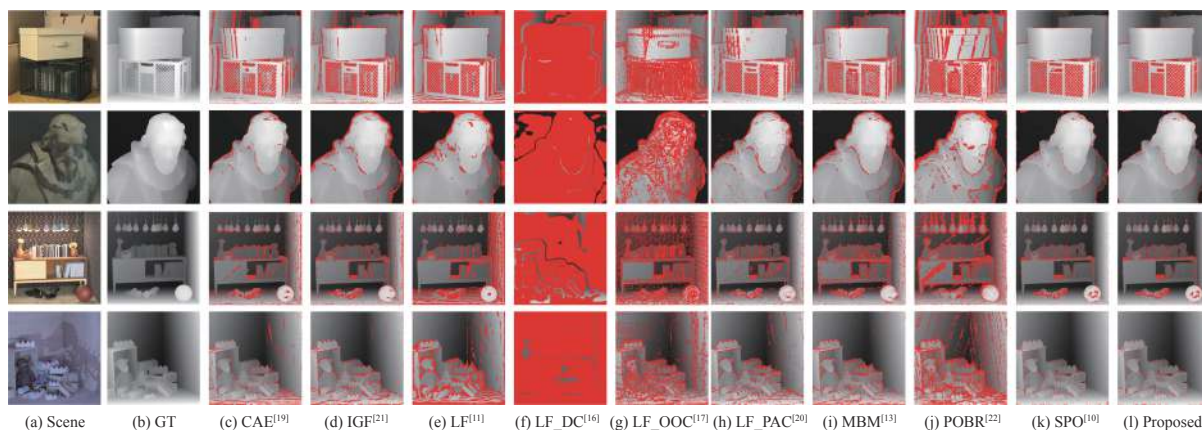


图 7 深度误点图结果比

Fig.7 Comparison of the results of the depth bad pixel map

如图 8 所示的 boxes 细节部分,文中提出的算法相比于其他算法在盒子的边缘区域处理的更好,对塑

料框的上沿也处理的很好; dino 细节部分的橱柜和积木边缘也较其他算法更为清晰准确。



在定量评估方面,文中采取了均方误差 MSE 和坏点率 BP 评判标准:

$$MSE = \frac{1}{N} \sum_p |GT(p) - \alpha^*(D)|^2 \quad (21)$$

$$BP = \frac{1}{N} \sum_p p |GT(p) - \alpha^*(p)| > \delta \quad (22)$$

式中:  $\alpha^*$  为  $p$  点计算出的深度;  $GT(p)$  为  $p$  点真值深度;  $N$  为像素点个数;  $\delta$  是深度误差阈值,如果计算结果中的深度值与深度真值差距大于阈值则会将该像素点判定为坏点,此次选择  $BP > 0.07$  作为评判标准。

定量指标上,如表 2、表 3 所示,文中所提出的算法在所有图片的  $BP > 0.07$  上,与其他算法相比达到了

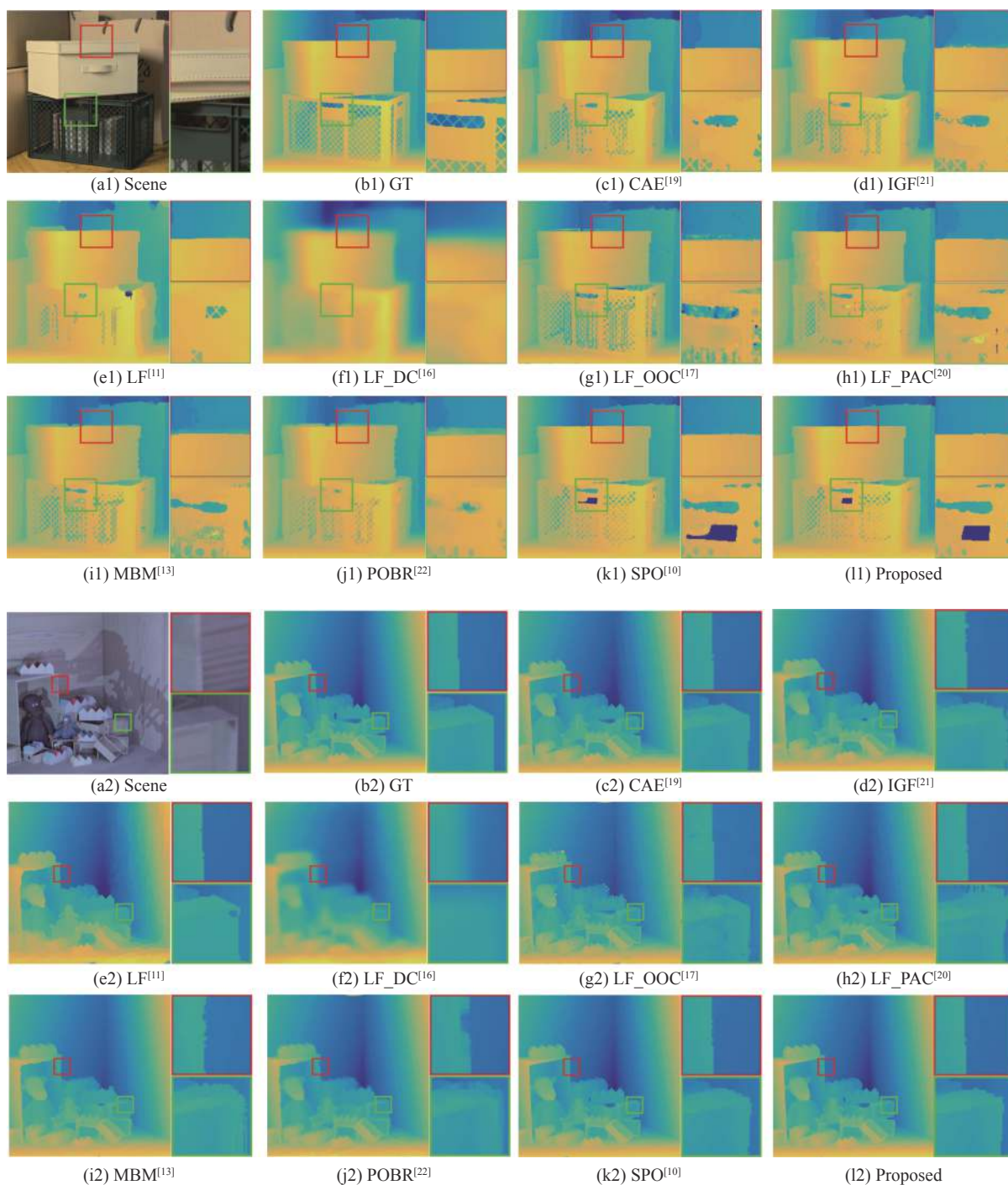


图 8 boxes, dino 深度图细节结果比较

Fig.8 Comparison of boxes and dino depth map detail results

表 2 BP>0.07 结果对比

Tab.2 Comparison of results of BP>0.07

	BP>0.07			
	boxes	Cotton	Sideboard	dino
CAE	0.187 8	0.038 4	0.120 0	0.068 8
IGF	0.192 4	0.141 8	0.108 3	0.070 5
LF	0.287 4	0.100 1	0.243 2	0.210 9
LF_DC	0.893 3	0.889 7	0.957 8	0.869 1
LF_OOC	0.405 6	0.365 5	0.436 4	0.252 7
LF_PAC	0.225 5	0.073 7	0.100 5	0.091 0
MBM	0.201 3	0.068 4	0.163 3	0.082 9
POBR	0.397 8	0.120 6	0.296 2	0.271 8
SPO	0.143 3	0.031 8	0.073 3	0.025 1
Proposed	0.141 1	0.025 5	0.071 9	0.023 7

表 3 MSE 结果对比

Tab.3 Comparison of results of MSE

	MSE			
	boxes	Cotton	Sideboard	dino
CAE	0.089 1	0.026 9	0.008 6	0.004 0
IGF	0.106 1	0.108 3	0.012 0	0.009 9
LF	0.162 9	0.141 0	0.039 0	0.017 1
LF_DC	0.121 2	0.061 8	0.575 5	0.030 5
LF_OOC	0.074 8	0.056 7	0.043 0	0.017 2
LF_PAC	0.099 2	0.070 5	0.011 0	0.009 2
MBM	0.101 9	0.067 1	0.029 1	0.013 7
POBR	0.134 8	0.063 5	0.046 8	0.023 4
SPO	0.108 8	0.041 4	0.009 9	0.003 9
Proposed	0.101 7	0.024 8	0.009 6	0.003 5

最佳, MSE 也有两张图片达到最佳, 综合效果也十分杰出。

## 4 结 论

文中从光场极平面图入手, 根据 EPI 斜线斜率与深度的对应关系设计了基于旋转线性算子 SLO 的 EPI 斜线像素一致性代价函数, 并与基于 SPO 的 EPI 斜线两侧区域差异性代价函数置信度加权结合, 得到了兼顾两种特性的融合代价函数, 最后使用了引导滤波和基于 MRF 的全局优化的后处理方法, 得到最终的深度图。并将文中的结果与经典的传统光场深度估计算法进行了比较, 实验结果表明, 在边缘处

理和噪声鲁棒性上都有很大提升, 综合效果达到最佳。

## 参考文献:

- [1] Ren Ng, Marc Levoy, Mathieu Brédif, et al. Light field photography with a handheld- plenoptic camera[D]. US: Stanford University, 2005.
- [2] Lytro. The lytro camera[EB/OL]. [2011-10-21] <https://www.lytro.com>.
- [3] Raytrix.3d light field camera technology[EB/OL]. [2013-05-05] [https:// www.raytrix. de](https://www.raytrix.de).
- [4] Wang Jiahua, Du Shaojun, Zhang Xuanzhe, et al. Design of focused light field computational imaging system with four-types focal lengths [J]. *Infrared and Laser Engineering*, 2019, 48(2): 0218003. (in Chinese)
- [5] Zhang Xuanzhe, Wang Yan, Wang Jiahua, et al. Image clarification and point cloud calculation under turbulence by light field camera [J]. *Infrared and Laser Engineering*, 2020, 49(11): 20200053. (in Chinese)
- [6] Bolles R C, Baker H H, Ma-rimont D H. Epipolar-plane image analysis: An approach to determining structure from motion [J]. *International Journal of Computer Vision*, 1987, 1(1): 7-55.
- [7] Jeon H G, Park J, Choe G, et al. Accurate depth map estimation from a lenslet light field camera[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1547-1555.
- [8] Chen C, Lin H, Yu Z, et al. Light field stereo matching using bilateral statistics of surface cameras[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1518-1525.
- [9] Zhang S, Sheng H, Yang D, et al. Micro-lens-based matching for scene recovery in lenslet cameras [J]. *IEEE Transactions on Image Processing*, 2017, 27(3): 1060-1075.
- [10] Fan Xiaoting, Li Yi, Luo Xiaowei, et al. Depth estimation based on light field structure characteristic and multiview matching [J]. *Infrared and Laser Engineering*, 2019, 48(5): 0524001.
- [11] Tao M W, Hadap S, Malik J, et al. Depth from combining defocus and correspondence using light-field cameras[C]// Proceedings of the IEEE International Conference on Computer Vision, 2013: 673-680.
- [12] Tao M W, Srinivasan P P, Malik J, et al. Depth from shading, defocus, and correspondence using light-field angular coherence[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1940-1948.
- [13] Wang T C, Efros A A, Ramamoorthi R. Occlusion-aware depth

- estimation using light-field cameras[C]//Proceedings of the IEEE International Conference on Computer Vision, 2015: 3487-3495.
- [14] Williem W, Park I K. Robust light field depth estimation for noisy scene with occlusion[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 4396-4404.
- [15] Park I K, Lee K M. Robust light field depth estimation using occlusion-noise aware data costs [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 40(10): 2484-2497.
- [16] Guo Z, Wu J, Chen X, et al. Accurate light field depth estimation using multi-orientation partial angular coherence [J]. *IEEE Access*, 2019, 7: 169123.
- [17] Sheng H, Zhang S, Cao X, et al. Geometric occlusion analysis in depth estimation using integral guided filter for light-field image [J]. *IEEE Transactions on Image Processing*, 2017, 26(12): 5758-5771.
- [18] Chen J, Hou J, Ni Y, et al. Accurate light field depth estimation with superpixel regularization over partially occluded regions [J]. *IEEE Transactions on Image Processing*, 2018, 27(10): 4889-4900.
- [19] Wanner S, Goldluecke B. Globally consistent depth labeling of 4D light fields[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition, 2012: 41-48.
- [20] Wanner S, Goldluecke B. Variational light field analysis for disparity estimation and super-resolution [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 36(3): 606-619.
- [21] Kim C, Zimmer H, Pritch Y, et al. Scene reconstruction from high spatio-angular resolution light fields [J]. *ACM Trans Graph*, 2013, 32(4): 1-73.
- [22] Huang Z, Lin C W, Shao H C, et al. Consistency constrained reconstruction of depth maps from epipolar plane images[C]//ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019: 2292-2296.
- [23] Zhang S, Sheng H, Li C, et al. Robust depth estimation for light field via spinning parallelogram operator [J]. *Computer Vision and Image Understanding*, 2016, 145: 148-159.
- [24] Sheng H, Zhao P, Zhang S, et al. Occlusion-aware depth estimation for light field using multi-orientation EPIs [J]. *Pattern Recognition*, 2018, 74: 587-599.
- [25] Li J, Jin X. EPI-neighborhood distribution based light field depth estimation[C]//ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020: 2003-2007.
- [26] Schilling H, Diebold M, Rother C, et al. Trust your model: Light field depth estimation with inline occlusion handling[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4530-4538.
- [27] Martin D R, Fowlkes C C, Malik J. Learning to detect natural image boundaries using local brightness, color, and texture cues [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(5): 530-549.
- [28] He K, Sun J, Tang X. Guided image filtering[C]//European Conference on Computer Vision, 2010: 1-14.
- [29] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(11): 1222-1239.
- [30] Honauer K, Johannsen O, Kondermann D, et al. A dataset and evaluation methodology for depth estimation on 4D light fields[C]//Asian Conference on Computer Vision, 2016: 19-34.