基于遮挡场景的光场图像深度估计方法

刘晓旻*,杜梦珠,马治邦,朱云飞,陈澎勃,马凤英

郑州大学物理学院,河南郑州 450001

摘要 针对在深度估计过程中的遮挡问题,提出一种新的基于多线索融合的光场图像深度估计方法。利用约束性 自适应散焦算法和约束性角熵度量算法获取场景的散焦线索、一致性线索,并计算出场景的初始深度、置信度。为 增强图像的边缘轮廓信息,通过 Canny 算子提取中心视角图像的边缘信息,然后利用马尔可夫随机场融合场景的 初始深度、置信度及边缘信息,实现图像的高精度深度估计。与其他先进方法相比,所提方法能够较好地解决场景 中存在的遮挡问题,获取的深度图精度较高、平滑效果较好,图像边缘保持效果较好。

关键词 图像处理;四维光场;深度估计;遮挡;多信息融合

中图分类号 TP391.4;O439 **文献标志码** A

doi: 10.3788/AOS202040.0510002

Depth Estimation Method of Light Field Image Based on Occlusion Scene

Liu Xiaomin^{*}, Du Mengzhu, Ma Zhibang, Zhu Yunfei, Chen Pengbo, Ma Fengying School of Physics, Zhengzhou University, Zhengzhou, Henan 450001, China

Abstract In this study, we develop a novel depth estimation method of light field image based on multi-cue fusion to solve the problem of occlusion in depth estimation process. Further, we employ the constrained adaptive defocus algorithm and the constrained angular entropy metric algorithm to obtain the defocusing and consistency clues of the scene, respectively. Subsequently, we determine the initial depth and confidence of the scene. The Canny operator is used to extract the edge information of the central perspective image for enhancing the image edge contour information. The initial depth, confidence, and edge information obtained from these scenes are fused using a Markov random field to achieve high-precision depth estimation. When compared with other advanced algorithms, the proposed method is more effective in solving the occlusion issue. The obtained depth map has higher precision, better smoothing effect and better edge retention effect.

Key words image processing; four-dimensional light field; depth estimation; occlusion; multi-information fusion OCIS codes 100.2000; 100.2960; 150.1488

1 引 言

目前,光场图像的深度估计方法主要有3类,即 基于极平面、基于重聚焦及基于多视角立体匹配的 深度估计方法。

基于极平面的方法是指将立体空间中的一点按 照不同的视角进行投影的方法,形成的投影点会形 成一条直线,该直线斜率与该空间点的深度相关^[1], 即利用极平面图(EPI)的结构张量来得到场景空间 点的深度信息。例如:Wanner 等^[2-3]利用 EPI 的结 构张量获取了局部深度信息,并通过可见性约束优 化得到了全局深度信息;Chen等^[4]利用 EPI 得到了 深度一致性的曲线,同时利用双边一致性度量方法 对图像匹配的可靠性进行了度量,从而实现图像的 深度估计。上述方法中,EPI 的计算过程涉及的数 据量较大而且时间成本较高,所以实时性较差,场景 的深度估计结果对遮挡较敏感。

基于重聚焦的方法是一种根据光场数字重聚焦 原理,对重聚焦后的图像进行聚焦性检测,从而估计 出深度信息的方法。例如,Tao 等^[5]利用散焦、相关

收稿日期: 2019-09-03; 修回日期: 2019-09-16; 录用日期: 2019-11-18

基金项目:国家自然科学基金(61307019)、河南省科技攻关重点项目(192102210304)、河南省科技开放合作计划项目 (152106000045)、河南省高等学校重点科研项目(18A140032)

^{*} E-mail: liuxmamara@zzu.edu.cn

性等线索提取焦点堆栈图像的局部深度图,并利用 马尔可夫随机场(MRF)对提取的线索进行融合,从 而获得最终的深度估计结果。这个方法仅适用于场 景简单且目标距离较近的情况,对于场景中存在遮 挡或者目标距离较远的情况,深度估计的结果会出 现较大的误差。Lu 等^[6]根据 LytroDesk^[7]获取重 聚焦图像,并利用 Laplace 算子检测对深度信息进 行了提取。这种方法获取的深度图中的场景纹理信 息较多,但会引入散焦误差,从而降低精确度。

基于多视角立体匹配的方法是一种利用光场图像的颜色一致性信息,通过构造成本量来提取深度信息的方法。Jeon 等^[8]针对光场深度估计中存在的窄基线问题,利用相移理论来获取微透镜阵列间的亚像素位移,然后将中心视角与其他视角分别进行立体匹配,最后利用图割优化算法获得最终的深度图。这个方法获取的深度图的精度较高,但需要对每个视角下的图像进行成本构造,因此耗时巨大,且对于遮挡问题,这种方法并不有效。

上述方法在一定场景中可获得较为准确的深度 图,但当场景存在遮挡时,深度估计的鲁棒性会明显 降低。为解决这一问题,本文在文献[9]的基础上提 出一种新的基于多线索融合的且对遮挡场景具有鲁 棒性的深度估计方法。利用约束性自适应散焦方法 描述场景的散焦信息,利用角熵信息衡量场景的一 致性信息计算出场景的初始深度、置信度,利用 Canny 算子对中心视角图像进行边缘性检测,从而 获得图像的边缘信息。然后通过 MRF 融合场景的 初始深度、置信度及边缘信息,通过最大后验概率马 尔可夫随机场(MAP-MRF)模型的最小化能量函数 对深度估计结果进行全局一致优化,最终实现图像 的高精度深度估计。

2 光场的成像原理与数字重聚焦

光场相机由主透镜、微透镜阵列及成像传感器 组合而成^[10],其结构如图1所示。物体上一点向各 个视角反射光线,光线经主透镜会聚后进入微透镜 阵列,然后按照入射方向分散到像素传感器上,因此 光场相机可以记录同一空间点发出的不同视角的光 线。图1中 uv 表示主透镜平面,xy 表示图像传感 器平面,d 表示微透镜直径,f 表示微透镜阵列与图 像传感器间的距离。

图 2 为包含物点空间信息与角度信息的四维(4D) 光场图像。当确定角度坐标 u 和在整个 xy 坐标空间 变化的角度坐标 v,就可得到多视角图像,如图 3 所示。









图 2 4D 光场示意图 Fig. 2 Diagram of 4D light field





利用光场图像的多视角特性可进行数字重聚 焦^[11],从而获得重聚焦图像,其原理如图4所示,其中 Π表示微透镜阵列平面,Ψ表示图像传感器平面,Φ表 示重聚焦平面,F表示透镜焦距,α表示聚焦系数(聚焦 深度),αF表示移动像面与微透镜阵列平面的距离。





(2)

重聚焦公式为

$$L_{aF}(x, y, u, v) = L_F\left[x + u\left(1 - \frac{1}{\alpha}\right), y + v\left(1 - \frac{1}{\alpha}\right), u, v\right], \qquad (1)$$

式中:LF 为输入的光场图像;LaF 为重聚焦图像。

获得多视角图像与重聚焦图像后,利用散焦线 索与一致性线索可估计出光场图像的深度信息。

3 本文方法

3.1 约束性自适应散焦算法

获取光场图像的深度信息时,传统的散焦算法 对遮挡不具有鲁棒性,因此可利用约束性自适应散 焦算法估计场景的散焦线索,获取场景的散焦深度 响应。当光场重聚焦在α处,角度像素块存在遮挡 时,可将角度像素块划分为多个大小相同的子斑块, 以获得不受遮挡影响的子斑块。每个子斑块的散焦 响应为

$$D_{c,a}(x,y) = \frac{1}{|N_{c}(x,y)|} \sum_{(x',y') \in N_{c}(x,y)} |R_{a}(x',y') - P(x',y')|,$$

式中:c 为子斑块的索引; $|N_c(x,y)|$ 为第c 个子斑 块的像素个数; $R_a(x',y')$ 为子斑块中(x',y')位置 处的像素值;P 为中心视角图像。自适应散焦响应 为子斑块的最小散焦响应,也表示子斑块 c^* 的散焦 响应,即 $c^* = \min_{c} D_{c,a}(x,y)$ 。为得到更加准确的 散焦信息,可加入颜色一致性约束 $D_{col,a}(x,y)$ 对结 果进行优化。传统的散焦算法是先计算子斑块的最 小散焦响应,而后加入约束条件。而本文在计算出 最小散焦响应之前加入 $D_{col,a}(x,y)$,将约束条件细 化到散焦响应中,以此减小散焦深度图中存在的噪 声。其效果对比如图 5 所示。



图 5 效果对比图。(a)约束性自适应散焦深度图; (b)传统散焦深度图



颜色一致性约束表达式为
$$D_{col,a}(x,y) = \min |R_a(x',y') - P(x,y)|.$$
(3)

(3)式可表述为子斑块像素与中心视角像素差的最小值。通过计算每一个子斑块的散焦响应和颜色一致性约束,最终能确定出最小散焦响应。所以约束性自适应散焦响应公式为

$$D_{\alpha}(x,y) = \min \left[D_{c,\alpha}(x,y) + \gamma D_{col,\alpha}(x,y) \right],$$
(4)

式中:γ为约束参数。角度像素块与子斑块的大小 可影响散焦响应的结果,因此在实验时要选择合适 尺寸的角度像素块和子斑块。

3.2 约束性角熵度量算法

场景的一致性信息是通过测量角度像素块的像 素相似性得到的。然而,传统的散焦算法未考虑场 景中存在遮挡的情况,使得深度估计结果不准确。 当场景存在遮挡时,角度像素块中受遮挡影响的像 素不满足一致性,但剩余部分像素依旧满足一致性, 因此提出利用优势像素的强度概率来描述该角度像 素块的一致性,并利用角度像素块的熵值信息得到 场景的一致性信息。假设光场重聚焦在 α 处,场景 无遮挡时,角度像素块中的像素满足一致性,此时像 素强度值集中,角度像素块的强度直方图为窄的高 峰,对应的熵值较低;当场景有遮挡时,像素强度值 较分散,强度直方图为宽的低峰,此时对应的熵值较 高。熵值越低,一致性信息越可靠,因此可利用角度 像素块的熵值响应来获得场景的一致性深度响应, 即利用角熵度量算法计算出角度像素块的熵值,从 而获得对遮挡具有稳鲁棒的一致性深度响应。熵值 的计算公式为

$$H_{\alpha}(x,y) = -\sum_{i} h(i) \log \left[h(i)\right], \qquad (5)$$

式中:*i* 为灰度;*h*(*i*)为角度像素块灰度*i* 的概率。

为降低遮挡对结果的影响,保证数据的稳定性, 可对角度像素块中的每一个像素强度赋予权重,即 引入权重函数 w(i),公式为

$$w(i) = \exp\left[-\frac{|i - A_{\alpha}(u_{c}, v_{c})|^{2}}{2\sigma^{2}}\right], \quad (6)$$

式中:(u_c,v_c)为像素中心视角坐标; σ为标准差。

当角度像素块中像素强度 *i* 与中心视角像素强 度 A_a(*u*_c,*v*_c)近似相等时,权重为 1;当远离中心视角 像素强度时,权重几乎为 0。因此,赋予权重的方法 可以降低遮挡对一致性结果的影响,此方法被称为约 束性法。此时角度像素块中像素强度 *i* 的概率为

$$g(i) = w(i)h(i)$$
。 (7)
由(5)、(7)式可知,约束性熵值的公式为

$$H_{a}(x,y) = -\sum_{i} \frac{g(i)}{|g|} \log [g(i)], \qquad (8)$$

式中:|g|为所有像素强度概率之和; $\frac{g(i)}{|g|}$ 为像素强度概率的归一化。

在计算过程中,分别计算 RGB 三个颜色通道的 熵值,为了整合三个通道的数据,提出两个方法,即 最大池化法和平均池化法,表达式分别为

$$C_{\max,a}(x,y) = \max[H_{R,a}(x,y), H_{G,a}(x,y), H_{B,a}(x,y)], \qquad (9)$$

$$C_{\text{avg, }a}(x,y) = \frac{H_{R,a}(x,y) + H_{G,a}(x,y) + H_{B,a}(x,y)}{3} \circ$$
(10)

当存在一个主导的颜色通道时,最大池化法可 获得更好的效果,例如,蓝色物体在蓝色通道中强度 高,在红色和绿色通道中强度近似为 0。其他情况 则平均池化法的效果较好。为了得到在不同成像条 件下的一致性信息,现将两种方法进行融合,最终的 熵值公式表示为

 $C_{\alpha}(x,y) = \beta C_{\max,\alpha}(x,y) + (1-\beta)C_{\alpha \vee g,\alpha}(x,y),$ (11)

式中:β为权重系数,β∈[0,1]。

3.3 初始深度和置信度

在获得重聚焦在 α 处的散焦信息 $D_a(x,y)$ 与 一致性信息 $C_a(x,y)$ 后,将二者进行求和运算,得 到场景的初始深度 Depth_a(x,y),其表达式为

 $Depth_a(x,y) = C_a(x,y) + D_a(x,y), \quad (12)$ $Depth(x,y) = \arg\min\{Depth_a(x,y)\}_{\circ} \quad (13)$

为表征初始深度估计的可靠性,用相应深度信息的置信度来表示,即利用同一像素在不同深度聚 焦面上的深度值进行对比,将最大深度值与次大深 度值的比值作为置信度(Depthconf),比值越大,该点 深度估计值可靠性越高。Depthconf表示为

 $\operatorname{Depth}_{\operatorname{conf}}(x,y) = \operatorname{Depth}_{a}(x,y) / \operatorname{Depth}_{a^{2}}(x,y)$

3.4 最终的深度估计算法

为增强图像的边缘轮廓信息,可利用 Canny 算 子对中心视角图像进行边缘检测,进而得到场景的 边缘信息*V_i*(*p*,*q*)。在进行最后的深度估计时,将 置信度作为场景的先验信息,以优化深度图的错误 区域。通过 MAP-MRF 模型融合场景的初始深度 信息、边缘检测信息,以获得最终的深度估计结果。 利用 MAP-MRF 模型的最小化能量函数,进行全局 一致的图像深度估计最优化。能量函数为

$$E = \sum_{p} E_{\text{unary}} [p, d(p)] + \sum_{p,q} E_{\text{binary}} [p, q, d(p), d(q)], \quad (15)$$

式中: *p*、*q*分别为空间中的两个相邻像素; *d*(*p*)、 *d*(*q*)分别为*p*、*q*两点对应的优化深度; *E*_{unary}为某 一空间点的散焦信息、一致性信息及置信度之间的 运算, 表征该点深度估计的可靠性; *E*_{binary}为相邻像 素的平滑约束。*E*_{binary}的运算可以表示为

$$E_{\text{binary}}[p,q,d(p),d(q)] = \frac{\exp\left[-d(p) - d(q)\right]^{2}}{|\nabla_{i}(p,q)|}$$
(16)

式中:*\nabla_i(p,q)*为中心视角图像中两相邻像素的灰度变化,即边缘信息。

4 实验结果与分析

4.1 实验设置

实验设备有北京卓立汉光微米级三维平移台 2 个,维视图像工业相机 MV-VEM200SC 1 个,分辨 率为 1200 pixel×1600 pixel。对实验数据进行处理 的软件为 MATLAB2016b 和 vs2010。计算机为 Intel(R) Core i7-3770,参数为 CPU 3.40 GHz、内存 16 GB、Windows7 64 位。

4.2 定性分析

为验证所提深度估计算法的准确性,首先对 HCI公开光场基准数据集 MonasRoom^[12]进行深度 估计。所提算法计算场景的散焦深度信息、一致性 深度信息结果如图 6 所示。

通过(13)、(14)式分别求出散焦信息与一致性



图 6 散焦深度图与一致性深度图。(a)约束性自适应散焦 深度图;(b)约束性角熵深度图

- Fig. 6 Defocus depth map and consistent depth map.
 - (a) Constrained adaptive defocus depth map;
 - (b) constrained angle entropy depth map

(14)

信息的初始深度、置信度,通过 MRF 融合场景的初始深度、置信度、边缘信息,并对深度估计结果进行

全局优化,得到最终的深度图。初始深度、置信度、 边缘检测的结果及最终深度图如图 7 所示。



图 7 图像结果。(a)初始深度图;(b)置信度图;(c)边缘检测结果;(d)最终深度估计结果

Fig. 7 Map results. (a) Initial depth map; (b) confidence map; (c) edge detection result; (d) final depth estimation result

为更加直观地展示最终深度估计的准确度,分 别利用 Wang 方法^[13]、Jeon 方法^[8]以及 Tao 方法^[5] 对 HCI 公开数据集^[12]进行深度估计,并与基准深度对比,其结果如图 8、9 所示。



图 8 MonasRoom 光场基准数据集的深度估计结果。(a) MonasRoom 中心视角图像;(b) GT 值;(c)所提方法; (d)文献[13]的方法;(e)文献[8]的方法;(f)文献[5]的方法

Fig. 8 Depth estimation results of MonasRoom light field reference dataset. (a) MonasRoom central perspective image;(b) GT value; (c) the proposed method; (d) method in Ref. [13]; (e) method in Ref. [8]; (f) method in Ref. [5]



图 9 Budden 光场基准数据集的深度估计结果。(a) Budden 中心视角图像;(b) GT 值;(c)所提方法; (d)文献[13]的方法;(e)文献[8]的方法;(f)文献[5]的方法

Fig. 9 Depth estimation results of Budden light field reference dataset. (a) Budden central perspective image;(b) GT value; (c) the proposed method; (d) method in Ref. [13]; (e) method in Ref. [8]; (f) method in Ref. [5]

由图 8、9 可知: Tao 的方法^[5]提取的深度图的 边缘模糊,颗粒噪声较多,分辨率低; Jeon 的方法^[8] 提取的深度图的颗粒噪声较少,且连续区域精确度 较高,但边缘区域过度平滑,存在错误估计区域; Wang 的方法^[13]和所提方法提取的深度图在遮挡区 域和边缘处的估计效果较好,但 Wang 的方法^[13]仅 适用于解决单一遮挡的情况,且易受噪声影响。针 对复杂的遮挡情况,所提方法的深度估计效果较好, 且几乎无颗粒噪声,边缘处信息更加准确,深度图的 平滑效果更好、更稳定。

图 10 为有遮挡和精细结构的真实场景的深度 估计结果。从图 10 可看出:所提方法能够正确估算 出物体形状,不会过度平滑物体边缘,边缘效果较 好;在遮挡区域,所提方法能够正确估算出深度信 息,不会存在边缘伪影,清晰度较高,图像平滑效果 较好。



图 10 Lytro 拍摄的 5×5 视角的光场图像的深度估计结果 Fig. 10 Depth estimation results of 5×5 view light field images taken by Lytro

4.3 定量分析

合成数据集提供了深度真值,因此可进行定量 分析。同文献[13],在合成光场数据集^[14]上对所提 方法进行深度估计实验,并进行定量分析。用方均 误差(MSE, *E*_{MS})作为定量指标来评价深度估计算 法的性能, *E*_{MS}越小,性能越好。*E*_{MS}的表达式为

$$E_{\rm MS} = \frac{1}{MN} \sum_{0 \le x \le M0 \le y \le N} \left[G_{\rm T}(x, y) - G(x, y) \right]^2,$$

(17)

式中: $G_{T}(x,y)$ 为(x,y)像素位置处的标准深度真

值;G(x,y)为(x,y)像素位置处的估计深度值;M、N 分别为深度图的尺寸。E_{MS}对比结果如表1所示。

Table 1 E_{MS} comparison

Method	MonasRoom	Budden	Medieval
Method in Ref.[13]	0.00338	0.00956	0.01272
Proposed method	0.00211	0.00413	0.01210

所提方法在遮挡区域、边缘处的深度估计效果较好,在丰富边缘信息上具有较大的优势。由表1可知,所提算法在各数据集上的 *E*_{MS}低于文献[13]的方

法,这与定性评价结果相一致,具有真实的可靠性。

5 结 论

针对光场图像深度估计过程中存在的遮挡问题,提出一种新的基于多线索融合的深度估计方法。 利用约束性自适应散焦方法和约束性角熵度量方法 获取场景的深度线索,计算出场景的初始深度和置 信度,这可有效降低遮挡对深度信息提取的影响;利 用 Canny 算子计算出中心视角图像的边缘信息,这 可有效提高深度图像边缘的清晰度;利用马尔可夫 随机场融合场景的初始深度、置信度及边缘信息,得 到全局一致的优化深度估计结果,可实现图像的高 精度深度估计。实验结果显示,所提方法获取的融 合深度图能有效减弱遮挡影响、减少噪声颗粒,使图 像变得更加平滑,边缘信息更加丰富。针对所提方 法,下一步将研究该方法对高光谱光场图像深度估 计的影响。

参考文献

- Xiong W, Zhang J, Gao X J, et al. Anti-occlusion light-field depth estimation from adaptive cost volume [J]. Journal of Image and Graphics, 2017, 22(12): 1709-1722. 熊伟,张骏,高欣健,等. 自适应成本量的抗遮挡光 场深度估计算法[J]. 中国图象图形学报, 2017, 22 (12): 1709-1722.
 Wanner S, Goldluecke B. Globallyconsistent depth
- [25] Walmer O, Golaldeene D. Gobalyconsistent depth labeling of 4D light fields[C] // 2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI, USA. New York: IEEE, 2012: 41-48.
- Wanner S, Goldluecke B. Variationallight field analysis for disparity estimation and super-resolution
 [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(3): 606-619.
- [4] Chen C, Lin HT, Yu Z, et al. Light field stereo matching using bilateral statistics of surface cameras
 [C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 1518-1525.
- [5] Tao M W, Hadap S, Malik J, et al. Depth from combining defocus and correspondence using light-field cameras [C] // 2013 IEEE International

Conference on Computer Vision, December 1-8, 2013, Sydney, Australia. New York: IEEE, 2013: 673-680.

- [6] Lu W, Mok W K, Neiman J, et al. 3D and image stitching with the Lytro light-field camera[D]. New York: City University of New York, 2013.
- [7] Cartpauj Software updates: Lytro illum firmware 2. 0, Lytro desktop 4.3[EB/OL]. [2019-09-03]. http: // lightfield-forum. com // tag // lytro-desktopsoftware.
- [8] Jeon H G, Park J, Choe G, et al. Accurate depth map estimation from a lenslet light field camera[C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 1547-1555.
- [9] Williem 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, 2018, 40(10): 2484-2497.
- [10] Sun F S, Han X. Multi-objective fast ranging method based on microlens array light field camera[J]. Acta Optica Sinica, 2019, 39(8): 0815001.
 孙福盛,韩燮.基于微透镜阵列型光场相机的多目标 快速 测 距 方 法 [J]. 光 学 学 报, 2019, 39(8): 0815001.
- [11] Liu R X, Ren M J, Xiao G B. A method of eliminating artifacts in focused light field imaging[J]. Acta Optica Sinica, 2018, 38(6): 0611001.
 刘润兴,任明俊,肖高博.一种聚焦式光场成像去混 叠方法[J].光学学报, 2018, 38(6): 0611001.
- Wanner S, Meister S, Goldluecke B. Datasets and benchmarks for densely sampled 4D light fields[C] // Annual Workshop on Vision, Modeling and Visualization(VMV), Lugano, Switzerland. [S.l.: s.n.], 2013: 225-226.
- [13] Wang T C, Efros A A, Ramamoorthi R. Depth estimation with occlusion modeling using light-field cameras[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(11): 2170-2181.
- Wang T C, Efros A A, Ramamoorthi R. Occlusion-aware depth estimation using light-field cameras [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 3487-3495.