

深度成像理论与实现

刘子伟,许廷发,王洪庆,申子宜,饶志涛

(北京理工大学 光电学院 光电成像技术与系统教育部重点实验室,北京 100081)

摘要:传统成像获取信息不足,成像质量有一定局限性。为此,提出了一种深度成像模型。模型包含深度矩阵、分解函数、散焦算子、自适应正则项等部分。深度矩阵的获取有双目立体视觉、结构光或飞行时间法等实现方法;分解函数用于将图像按深度值的不同分割为若干子图像;散焦算子可以通过深度散焦法来计算;自适应正则项的引入能减少图像的阶梯效应,增强图像的光滑性。通过局部标准差和局部平均梯度这两个评价指标检验深度成像模型的效果。实验结果表明,深度成像模型效果显著。

关键词:深度成像; 计算成像; 深度; 深度散焦法

中图分类号: TP391 **文献标志码:** A **DOI:** 10.3788/IRLA201645.0726001

Theory and implementation of depth photography

Liu Ziwei, Xu Tingfa, Wang Hongqing, Shen Ziyi, Rao Zhitao

(Key Laboratory of Photoelectronic Imaging Technology and System, Ministry of Education of China, School of Optoelectronics, Beijing Institute of Technology, Beijing 100081, China)

Abstract: Traditional imaging methods are subject to limitations on information acquisition, which brings about the deficiency of image quality. Therefore, a depth photography model was proposed here. This model contains a depth matrix, a decomposition function, a defocus operator, and an adaptive regularization term. The depth matrix could be estimated using the binocular stereo vision method, the structured light approach, or the time-of-flight algorithm. The decomposition function was used to segment the image into pieces according to the distinct depth values. The defocus operator was calculated through a defocus-from-depth method. The adaptive regularization term reduces the staircase effect and enhances image smoothness. Local standard deviations and local average gradients were used to evaluate the effectiveness of depth photography model. Experimental results demonstrate that the proposed model is effective.

Key words: depth photography; computational imaging; depth; defocus from depth

收稿日期:2015-11-24; 修订日期:2015-12-27

基金项目:国家自然科学基金(61172178, 61371132, 61471043);高等学校博士学科点专项科研基金(20121101110022);
国际科技项目(2014DFR10960)

作者简介:刘子伟(1987-),男,博士,主要研究方向为三维成像、计算成像。Email:lzwjmo@163.com

通讯作者:许廷发(1968-),男,教授,博士,主要研究方向为三维成像、计算成像以及光电成像探测与识别。Email:xutingfa@163.com

0 引言

成像技术的发展经历了四个阶段:小孔成像、透镜成像、数字成像、计算成像。计算成像有助于记录更丰富的视觉经历,捕捉更多的、像素之外的信息,获得机器可读性更强的情景描述。传统的成像方法由于捕获的信息不充分,在景深、动态范围和图像分辨率等方面有较明显的局限性。相比之下,计算成像由于结合了传统光学系统和新的光学结构以及附加了计算能力,能够获取更丰富的视觉信息,从而得到更高质量的图像。计算成像的主要研究领域包括光学掩模辅助^[1]、孔径改造^[2]、光路及传感器设计^[3-4]、集成成像^[5]、光场成像^[6]以及复合计算^[7-8]。

文中提出了一种基于深度的计算成像模型,其目标是扩展景深,并获得全对焦图像。给出了深度图(或深度矩阵)、分解函数、散焦图(或散焦矩阵)的计算公式和实现细节。

1 深度成像模型

1.1 深度检测方法

通常图像中的物体具有不同的清晰度,这与它们不同的深度值有关。对于图像来说,深度信息具有重要作用,有必要进行准确的检测。

双目立体视觉系统^[9]是获取深度信息的一种传统方法。它一般由一对相同的单目成像结构组成,能提供同一场景的两个不同的视场。另外,双目系统也可以由一个单目成像结构先后放置在两个不同的位置来实现。双目成像分为横向模式和轴向模式,这两种模式都是适用于几十米以内的小深度的检测,比如室内环境。

结构光^[10]利用了照明中的几何信息,能够作为大深度检测方法。一个结构光系统由一个相机、一个光源和景物组成。从光源发出的光抵达物体的表面,相机捕捉到景物被照射的部分,借助于三角计算可以获得深度信息。

飞行时间法^[11],根据激光的状态一般可分为脉冲激光测距或连续波激光测距。这种方法通过测量光从光源到景物并最终到达传感器需要的时间来计算深度。飞行时间法精度高,抗干扰能力强,非常适用于大深度的测量。

1.2 分解函数

在场景的深度图中,不同深度值的数量通常是有限且可数的。因此,可以设置一个计数变量通过遍历深度图来统计这个数值。假设原始的部分对焦的图像 g 具有 k 个不同的深度值,或其深度图有 k 个不同的值。图像 g 可以被分解为 k 个子图像,每个子图像具有相同的深度值,即 $g = \sum_{i=1}^{i=k} g_i$ 。或者

$$g_i = g \cdot r_i(z, i), i=1, 2, \dots, k \quad (1)$$

定义分解函数如公式(2)所示:

$$D_i = r_i(z, i) = I - R_d[|z - d_i| / (|z - d_i| + \alpha)], \quad i=1, 2, \dots, k \quad (2)$$

式中: I 为与深度矩阵具有相同尺度的值全为1的矩阵; α 为一个很小的正数,用来避免分母出现0的情况; R_d 为四舍五入函数;运算符./代表矩阵之间的点除运算。

根据公式(2)可以获得第 i 个分解因子 D_i ,其元素为1或0。 $D_i(x, y)=1$ 表示图像 g 中的像素 (x, y) 对应 D_i ,而 $D_i(x, y)=0$ 表示不对应第 i 个分解因子。

1.3 深度散焦法

通过计算散焦模糊的大小可以得到物体的距离(深度),即散焦深度法。与此相反,采用深度散焦法来获得散焦图。

1.3.1 散焦的类型

散焦可以被分为图像散焦、传感器散焦和场景散焦。图像散焦指图像处理领域中常用的散焦。

焦平面上物体的一点所发出的光会聚于传感器平面上的一个点;而散焦平面(物体上的点到透镜的距离不等于焦平面距离)上物体的一点所发出的光会在传感器平面上形成一个直径为 e 的圆,即传感器散焦。

另一方面,如果一个物体不在焦平面上,传感器上的一个点所会聚的光线来自于散焦平面上的一个直径为 c 的圆形区域,即场景散焦。图1为传感器散焦和场景散焦的几何原理。如图,当物体位于焦平面上,距离为 u_1 ,对应的清晰图像应位于 v_1 ;当物体位于散焦平面上,距离为 u_2 ,对应的清晰图像应出现在距离为 v_2 的位置。

传感器散焦 e 可以按如下公式计算:

$$e = \frac{|v_1 - v_2|}{v_2} \cdot D \quad (3)$$

式中： D 为孔径直径。场景散焦 c 可以表示为：

$$c = \frac{|u_2 - u_1|}{u_1} \cdot D \quad (4)$$

传感器散焦和场景散焦的关系可以由以上两个公式推导出：

$$\frac{e}{c} = \frac{|v_1 - v_2|}{v_2} \cdot \frac{u_1}{|u_2 - u_1|} \quad (5)$$

根据公式(6)可以获得图像散焦：

$$s = \frac{e}{h} \quad (6)$$

式中： h 为每像素对应的宽度。假设图像尺寸为 $m \times n$ (像素 \times 像素)，传感器尺寸为 $p \times q$ (mm \times mm)，则 $h = p/n = q/n$ (mm/pixel)。

公式(6)表明图像散焦正比于传感器散焦。以上就是一个成像系统中的三种散焦之间的主要关系。

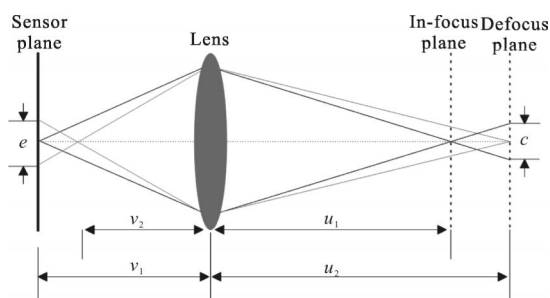


图 1 散焦形成原理

Fig.1 Principle for defocus formation

1.3.2 散焦的计算

首先考虑景物到传感器平面的距离已知的情况。在一个简化的场景中只包含两个物体，一个近景和一个远景，与传感器平面的距离分别为 z_1 和 z_2 。假设近景对焦，可以得到公式(7)~(10)：

$$v_1 + u_1 = z_1 \quad (7)$$

$$v_1 + u_2 = z \quad (8)$$

$$\frac{1}{v_1} + \frac{1}{u_1} = \frac{1}{f} \quad (9)$$

$$\frac{1}{v_2} + \frac{1}{u_2} = \frac{1}{f} \quad (10)$$

式中： f 为焦距。由于近景对焦，其散焦值为 0。根据公式(7)~(10)可以得到 v_1 和 v_2 的计算方法：

$$v_1 = \frac{z_1 - \sqrt{z_1^2 - 4z_1f}}{2} \quad (11)$$

$$v_2 = \frac{(2z - z_1 + \sqrt{z_1^2 - 4z_1f})f}{2z - z_1 + \sqrt{z_1^2 - 4z_1f} - 2f} \quad (12)$$

通过公式(3)、(11)和(12)可以推导出远景的传感器散焦，进而得到图像散焦 s ：

$$s = \left(\frac{z_1 - \sqrt{z_1^2 - 4z_1f}}{2f} - \frac{2z}{2z - z_1 + \sqrt{z_1^2 - 4z_1f}} \right) \cdot \frac{D}{h} \quad (13)$$

在公式(13)中，因为假定近景对焦，所以 z_1 为一个常量，并且 f 、 D 、 h 在一幅图像中也是不变的。因此，可以简化公式(13)为：

$$s(z) = k \left(t - \frac{z}{z - ft} \right) \quad (14)$$

式中： $k = D/h$ ， $t = (z_1 - \sqrt{z_1^2 - 4z_1f})/(2f)$ 。显然， $s(z_1) = 0$ 。

公式(15)即散焦模糊模型 $r_2(s)$ ：

$$r_2(s(x,y)) = r_2(x,y,s) = \begin{cases} \frac{4}{\pi s^2}, & x^2 + y^2 \leq \frac{s^2}{4} \\ 0, & \text{其他} \end{cases} \quad (15)$$

1.4 成像模型和实现算法

成像模型可以表示为：

$$f = \sum_{i=1}^k \operatorname{argmin}_f \int_{\Omega} |g \cdot r_1(z,i) - f_i^* r_2(s)|^2 + \lambda |\nabla f_i|^{1+(|\nabla f_i|^2)^{-1}} \cdot [1+(1+|\nabla f_i|^2)^{-1}]^{-1} d\Omega \quad (16)$$

公式(16)的核心部分为：

$$E(f_i) = \int_{\Omega} |g \cdot r_1(z,i) - f_i^* r_2(s)|^2 + \lambda |\nabla f_i|^{1+(|\nabla f_i|^2)^{-1}} \cdot [1+(1+|\nabla f_i|^2)^{-1}]^{-1} d\Omega \quad (17)$$

假设 $p(|\nabla f_i|) = 1+(1+|\nabla f_i|^2)^{-1}$ 。公式(17)由一个能量函数(第一项，用于计算最清晰的图像)和一个自适应项(第二项，能有效地减少图像的阶梯效应，增强图像的光滑性)构成。公式中的指数 $p(|\nabla f_i|) = 1+(1+|\nabla f_i|^2)^{-1}$ 取决于梯度的大小 $|\nabla f_i|$ 并且符合两个原则：单调递减；如果 $|\nabla f_i| \rightarrow 0$ ，那么 $p(|\nabla f_i|) \rightarrow 2$ ，如果 $|\nabla f_i| \rightarrow \infty$ ，则 $p(|\nabla f_i|) \rightarrow 1$ 。为简化运算，采用 $|\nabla G_{\sigma} * f_i|$ (G_{σ} 表示高斯滤波)代替 $|\nabla f_i|$ ，当 f_i 随迭代而改变时，这个指数就不用重新计算了。

利用变分法推导出公式(17)的欧拉-拉格朗日方程：

$$[g \cdot r_1(z,i) - f_i^* r_2(s)] * r_2(s) + \lambda \operatorname{div}(|\nabla f_i|^{p-2} |\nabla f_i|) = 0 \quad (18)$$

采用最速下降流算法计算公式(17)的最小值：

$$\frac{\partial f_i}{\partial t} = [f_i^* r_2(s) - g \cdot r_1(z,i)] * r_2(s) - \lambda \operatorname{div}(|\nabla f_i|^{p-2} |\nabla f_i|) \quad (19)$$

2 实验

通过一个真实场景的实验来验证文中提出的成像模型的有效性。实验场景主要包括 4 个桌子挡板和一面墙。因此,5 个分解因子被用到,即 D_1 、 D_2 、 D_3 、 D_4 和 D_5 。实验采用佳能 EOS-60D 相机,相机采用光圈优先模式,焦距为 1.8 cm,孔径为 $F/3.5$,距离最近的挡板设为对焦。

图 2 为传统成像和文中全对焦成像的结果。

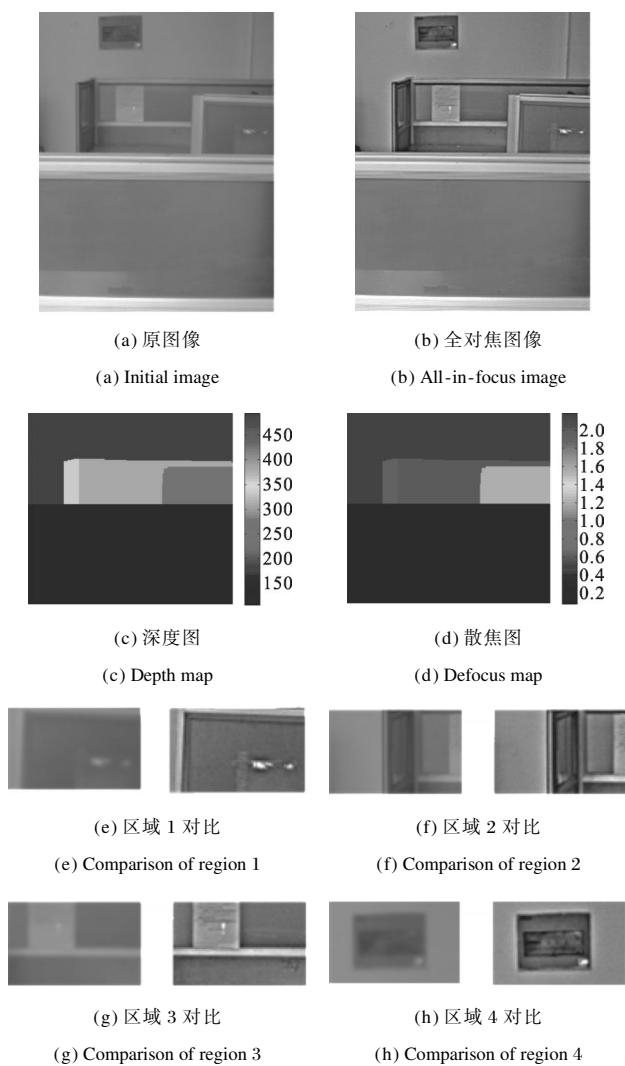


图 2 真实场景成像示例

Fig.2 Photography results of an actual scene

采用局部标准差(公式(20))和局部平均梯度(公式(21))来评价原始图像和全对焦图像的质量。图像被分割成 5×5 的小块,并分别计算这两个指标。

$$sd = \left[\frac{1}{n-1} \sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{i=1}^n x_i \right)^2 \right]^{1/2} \quad (20)$$

$$ag = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \left[\frac{1}{2} \left(\frac{\partial f_{i,j}}{\partial x_i} \right)^2 + \frac{1}{2} \left(\frac{\partial f_{i,j}}{\partial y_j} \right)^2 \right]^{1/2} \quad (21)$$

图 3 为局部标准差和局部平均梯度。通过对比可以看出,有一部分图像块的两个参数几乎重合,这是因为相应的图像块来自于场景中最近的挡板,而这个挡板是处于对焦状态。

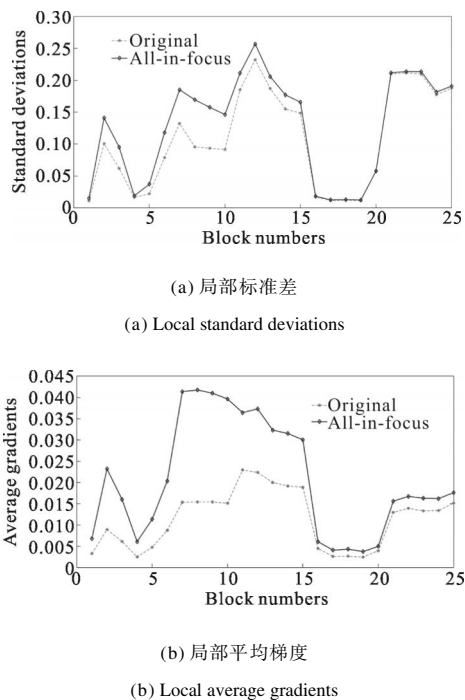


图 3 局部标准差和局部平均梯度

Fig.3 Local standard deviations and local average gradients

3 结论

文中提出了一种基于深度的计算成像模型,以获得全对焦图像。利用双目立体视觉、结构光或飞行时间法来计算场景的深度图。模型由深度矩阵、分解函数、散焦模糊算子以及一个自适应正则项构成。深度成像模型丰富了计算成像理论。

参考文献:

[1] Lv Qinghua, Zhu Xiao, Xu Zhongbao, et al. Extended depth of field with the optimal '0, π ' binary phase pupil mask[J]. *Journal of Modern Optics*, 2013, 60(7): 569-577.

[2] Shu Xianbiao, Gao Chunyu, Narendra A. Aperture access and manipulation for computational imaging[J]. *Computer Vision and Image Understanding*, 2012, 116(2): 222-237.

[3] Ma Xianglu, Feng Ying, Cao Yu. Concave hyperboloidal catadioptric omnidirectional imaging system[J]. *Infrared and*

- Laser Engineering*, 2013, 42(8): 2132–2136. (in Chinese)
马相路, 冯莹, 曹毓. 双曲凹面折反射全景成像系统[J]. 红外与激光工程, 2013, 42(8): 2132–2136.
- [4] Xiao Longlong, Liu Kun, Han Dapeng, et al. Focal plane coding method for high resolution infrared imaging [J]. *Infrared and Laser Engineering*, 2011, 40(11): 2065–2070. (in Chinese)
肖龙龙, 刘昆, 韩大鹏, 等. 焦平面编码高分辨率红外成像方法[J]. 红外与激光工程, 2011, 40(11): 2065–2070.
- [5] Xiao X, Bahram J, Genaro S, et al. Three-dimensional polarimetric computational integral imaging[J]. *Optics Express*, 2012, 20(14): 15481–15488.
- [6] Tomoya N, Ryoichi H, Jun T. Computational phase modulation in light field imaging [J]. *Optics Express*, 2013, 21(24): 29523–29543.
- [7] Jorge O C, Emmanuel Y V, Cristina M G S. Multiple-frame photography for extended depth of field [J]. *Applied Optics*, 2013, 52(10): D84–D91.
- [8] Bai Jian, Huang Zhi, Hou Xiyun. Design of panoramic stereo imaging with single optical system [J]. *Infrared and Laser Engineering*, 2012, 41(12): 3342–3348. (in Chinese)
白剑, 黄治, 侯西云. 单光学系统全景环带立体成像技术[J]. 红外与激光工程, 2012, 41(12): 3342–3348.
- [9] Wan Yilong, Bai Lianfa, Han Jing, et al. Method and realization of significant target recognition and distance ranging in the binocular stereo vision under low illumination [J]. *Infrared and Laser Engineering*, 2015, 44 (3): 1053–1060. (in Chinese)
万一龙, 柏连发, 韩静, 等. 低照度双目立体显著目标距离测定方法与实现[J]. 红外与激光工程, 2015, 44(3): 1053–1060.
- [10] Daniel S, Richard S. High-accuracy stereo depth maps using structured light [C]//Proceedings of the Computer Vision and Pattern Recognition, 2003: 195–202.
- [11] Jiejie Z, Liang W, Ruigang Y, et al. Fusion of time-of-flight depth and stereo for high accuracy depth maps [C]//Proceedings of the Computer Vision and Pattern Recognition, 2008: 3262–3269.