

文章编号:1002-2082(2019)02-0179-07

压缩光场重建与深度估计

刘晓旻,马治邦,王前程,杜梦珠,朱云飞,马凤英,梁二军

(郑州大学 物理工程学院,河南 郑州 450001)

摘要:针对光场深度估计过程中数据量大、边缘处深度估计结果不准确问题,利用压缩感知原理重建光场,提出一种新的多信息融合的光场图像深度估计算法。利用压缩感知重建算法重建 5×5 视角光场数据,获取光场数据后首先移动子孔径实现重聚焦,然后利用角度像素块散焦线索和匹配线索计算出场景初始深度和置信度。计算图像边缘信息,通过融合初始深度、置信度、边缘信息获取最终深度。实现压缩光场仿真重建,并对仿真光场数据和公开光场数据进行深度估计,实验结果表明:可以仿真重建出 5×5 视角光场数据,且仿真重建的光场可用于深度估计。该深度估计算法在场景边缘处的深度估计结果边界清晰,层次分明,验证了重建光场深度估计的可行性与准确性。

关键词:应用光学;深度估计;压缩感知;光场重建

中图分类号:TN201;TP391.4;O439

文献标志码:A

DOI:10.5768/JAO201940.0201001

Compression light field reconstruction and depth estimation

LIU Xiaomin, MA Zhibang, WANG Qiancheng, DU Mengzhu, ZHU Yunfei, MA Fengying, LIANG Erjun
(School of Physics and Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: Aiming at the problem of large amount of the light field data and inaccurate depth estimation at the edge, the light field was reconstructed by compressive sensing principle and a depth estimation method based on multiple information fusion was proposed. Firstly, a 5×5 view light-field was reconstructed according to the compressed sensing reconstruction algorithm, and the refocusing was realized by moving the sub-aperture after obtaining the field data. Then, the initial depth and confidence of the scene were calculated by using the defocused response cues and correspondence response of angular patch. Finally, the image edge information was calculated, and the final depth was obtained by fusing the initial depth, confidence and edge information. The compressed light field simulation and physical reconstruction were realized, and the depth estimation of the simulated reconstruction light field and the open optical field data were also achieved. The experimental results show that the method proposed can simulate and reconstruct the 5×5 view light field data, which can be used for depth estimation. The depth estimation results of the proposed method at the edge of the scene is clear and hierarchical, which verifies the feasibility and accuracy of the depth estimation for the reconstructed light field based on compressed sensing theory.

Key words: applied optics; depth estimation; compressed sensing; light field reconstruction

收稿日期:2018-08-15; 修回日期:2018-10-09

基金项目:国家自然科学基金(61307019,11404290,11574276);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)

作者简介:刘晓旻(1973—),女,博士,副教授,硕士生导师,主要从事光谱图像三维重建、光场三维重建、压缩光场重建方面的研究。E-mail:liuxmamara@zzu.edu.cn

引言

光场是空间中每一个点和每一个方向辐射函数的总和^[1],光场图像蕴含了物方全部的三维信息,可用在重聚焦^[2]、超分辨^[3]、深度估计^[4]、显著性检测^[5]、以及三维重建^[6]等领域。目前获取光场的途径主要有 3 种:相机阵列^[7]或者单个相机进行移动^[1]、微透镜阵列法^[2]以及掩膜法^[8]获取光场。相机阵列成本高,体积庞大,限制了它的应用,单个相机移动采集光场仅限静态场景。微透镜阵列法采集光场虽然光场采集设备轻便,但该方法采集的光场的空间分辨率(光场子图像大小)和角度分辨率(光场视角个数)存在制约,不能得到和传感器分辨率一样的光场子图像。掩膜法是将掩膜加入到成像系统中,对光线进行调制,再通过重建算法恢复出光场数据,这种方法可以解决光场空间分辨率与角度分辨率制约问题,恢复出的光场子图像大小和采集到的光场图像大小一致。目前研究光场重建主要有 Ashok、Babacan 等人分别用光学编码方法采集压缩光场^[9-10],常雷等人分析掩膜光场成像的特点并仿真验证^[11],刘永春等人仿真验证基于掩膜的光场采集与重建^[12]。覃亚丽利用低秩矩阵分解完成光场稀疏采样和重建^[13]。光场的深度估计是指将采集的光场通过一系列计算获取场景的深度信息,是当今国内外研究热点之一。目前研究光场深度估计的主要有:Wanner 和 Kim 等人利用对极平面图(EPI)通过计算优化得到最终深度图^[4,14],但 EPI 计算深度计算量大,容易受噪声影响。Chen、Heber、Jeon 等通过求解不同微透镜所对应子图像间的匹配关系来获取深度信息^[15-17],对匹配噪声难以抑制。Tao 和 Wang 等人利用图像的散焦信息为基础,融合图像其他信息后得到最终深度图^[18-19],其深度估计结果边缘效果不理想。

综上所述,为了有效降低光场采集、处理和传输的数据量,提高深度估计结果边缘处的准确性,本文利用压缩感知原理仿真重建了压缩光场,同时在文献[19]的基础上提出了一种新的融合场景多种信息的深度估计算法,并实现对仿真重建出的光场的深度估计。通过 K-SVD 算法训练学习获取光场过完备字典,使得光场可以稀疏表示。

然后利用压缩感知重建算法仿真重建出 5×5 视角光场。获取光场数据后,实现光场的重聚焦,对于存在遮挡区域,采用方差较小的部分代替整个角度像素块,计算聚焦不同深度处的角度像素块的均值和方差之和的最小值作为初始深度,最大值和次最大值的比值作为置信度。然后计算中心视角图像边缘信息,最后通过融合初始深度、置信度、边缘信息得到最终深度。实验通过压缩感知原理仿真实现了多视角光场重建,并且对仿真重建光场和公开光场数据进行了深度估计,实验结果表明,本文的深度估计算法在场景边缘处的深度估计结果边界清晰,层次分明。

1 压缩感知原理

压缩感知是一种新的采样理论^[20],突破了奈奎斯特采样定律的限制,在采样频率低于信号最高频率的 2 倍时,也能精确重建信号,只要满足:1) 原始信号可以在稀疏基下稀疏表示;2) 稀疏基 Ψ 与测量矩阵 Φ 满足约束等距性(RIP)就能够高概率地重构原始信号。信号 $f (f \in R^n)$ 压缩感知理论采样过程可表示为

$$y = \Phi f \quad (1)$$

式中: $\Phi (\Phi \in R^{m \times n} (m < n))$ 是测量矩阵^[1]; y 是测量值。若信号 f 是稀疏的,则可通过 y 与 Φ 重建原始信号。但在自然界中信号往往是不稀疏的,因此需要对信号稀疏表示,对信号 $f (f \in R^n)$ 若存在稀疏基 $\Psi (\Psi \in R^{n \times d})$ 和稀疏向量 α ,则信号 f 可稀疏表示为

$$f = \Psi \alpha \quad (2)$$

那么压缩感知的测量公式可表示为

$$y = \Phi \Psi \alpha = \Theta \alpha \quad (3)$$

1.1 光场稀疏表示

由于光场的复杂性,光场不能由单一完备基来稀疏表示。本文选择过完备字典作为稀疏基 Ψ ,提高复杂光场稀疏表示的通用性。通过采集多组光场样本数据作为光场样本集,将光场样本集随机分成多个 $p \times q$ 光场碎片,再利用 K-SVD^[21] 算法训练学习获取这个稀疏基 $\Psi^{n \times d} (n < d)$,即光场过完备字典。图 1 是部分光场过完备字典可视化,字典冗余度为 1.6。右侧是 2 个放大的原子,

可以看到其明显的角度方向性。

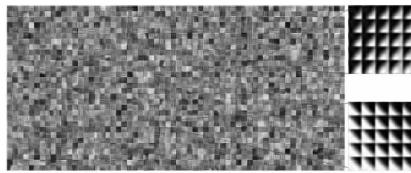


图 1 部分光场过完备字典可视化

Fig. 1 Part of visualization over-complete light field dictionary

1.2 光场测量矩阵设计

测量矩阵 Φ 将完整光场值传递到测量值 y , 完成信号的线性测量和降维, 即在采集光场图像时实现数据压缩。测量矩阵 Φ 与稀疏基 Ψ 共同构成传感矩阵 Θ , 重建原始信号时需要传感矩阵能够高概率地满足约束等距性^[21]。测量矩阵 Φ 与稀疏基 Ψ 之间高度不相干, 则传感矩阵 Θ 满足 k 阶约束等距性。为了最大程度地满足 Φ 与 Ψ 的不相干性, 更好重建 5×5 视角的 4D 光场图像, 对测量矩阵进行改进, 我们所用测量矩阵仍是随机矩阵, 但不直接生成等同图像空间分辨率为 808×808 像素的随机矩阵(如图 2 所示), 而是每次只生成等同分割块大小 8×8 像素的随机矩阵, 之后用多次生成的 8×8 像素大小随机矩阵组合成大小为 808×808 像素的矩阵作为测量矩阵(如图 3 所示)。这样的测量矩阵具有更高的随机性, 且与冗余字典能满足约束等距性(RIP)。

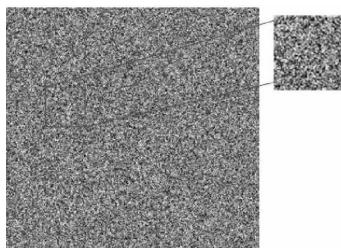


图 2 随机测量矩阵

Fig. 2 Random measurement matrix

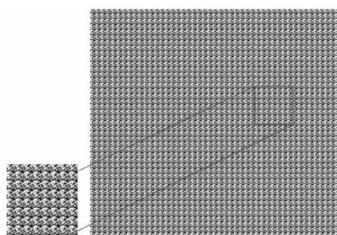


图 3 优化随机测量矩阵

Fig. 3 Optimization of random

1.3 光场重建算法

当传感矩阵满足阶约束等距性时, 用 l_1 最小化代替 l_0 最小化, 稀疏重建模型^[20]可以表示为

$$\min_a \|a\|_1 \quad s.t. \quad y = \Theta a \quad (4)$$

采用 LASSO (least absolute shrinkage and selection operator) 模型来求解。引入噪声约束参数 λ 更准确描述, 拉格朗日方程表示约束项形式为

$$\min_a \|y - \Theta a\|_2 + \lambda \|a\|_1 \quad (5)$$

采用交替方向乘子 ADMM^[22] (alternating direction method of multipliers) 算法求出稀疏向量 a , 再由过完备字典 Ψ 与稀疏向量求出原始光场。

重建时, 需要将采集的二维光场图像, 即测量值 y 转变为 $N \times 1$ 数组进行计算。计算时数据量大, 为了提高计算效率, 采用滑动分块法重建光场, 滑动模块大小为 8×8 像素。

2 光场深度估计原理

光场成像中, 可以通过计算像点角度像素块^[19] 特征来表征场景中的深度信息。通过数字重聚焦原理^[23] 获得不同深度聚焦序列图像。计算不同深度聚焦图像角度像素块的均值和方差来表征匹配线索和散焦线索。为使初始深度估计准确, 用匹配线索和散焦线索之和的最小值来表示初始深度。用匹配线索和散焦线索之和的最大值和次最大值比值表示置信度^[19]。通过融合场景初始深度、置信度、中心视角图像边缘信息来获取最终深度。

2.1 初始深度和置信度

当重聚焦到 α 处时。遮挡区域被遮挡物分割成 2 个区域, 方差小的区域用 i 表示, 匹配线索用 $C_a(x, y)$ 表示、散焦线索用 $D_a(x, y)$ 表示, 则有:

$$C_a(x, y) = V_{a,i}(x, y) \quad (6)$$

$$D_a(x, y) = (I_{a,i}(x, y) - L(0, 0, x, y))^2 \quad (7)$$

式中 $V_{a,i}(x, y)$ 、 $I_{a,i}(x, y)$ 分别表示聚焦在 α 处时, 角度像素片的方差和均值。

初始深度用 $Z(x, y)$ 表示, 置信度用 $W(x, y)$ 来表示, 有关系式:

$$H_a(x, y) = C_a(x, y) + D_a(x, y) \quad (8)$$

$$Z(x, y) = \arg \min_a H_a(x, y) \quad (9)$$

$$W(x, y) = H_a(x, y) / H_a^2(x, y) \quad (10)$$

2.2 最终深度计算

本文在最终深度的计算过程中融合图像边缘

信息来增加场景边缘处深度估计准确性,其最终深度的计算是利用马尔卡夫随机场来传播,定义能量函数为

$$E = \sum_p E_{\text{unary}}(p, W(p)) + \lambda \sum_{p,q} E_{\text{binary}}(p, q, W(p), W(q)) \quad (11)$$

E_{unary} 部分的计算类似于 Tao^[24] 里面 (MRF propagation) 表示初始深度 $Z(x, y)$ 与置信图 $W(x, y)$ 之间运算, 表征着该点深度估计可靠性。

中心视角图像相邻像素之间灰度变化可表示为 $\nabla i(p, q)$, 有关系式:

$$E_{\text{binary}}(p, q, W(p), W(q)) = \frac{\exp[-(W(p) - W(q))^2]}{|\nabla i(p, q)|} \quad (12)$$

式中: $W(p)$ 表示初始深度响应; p, q 表示两相邻像素; 分子部分表示相邻像素初始深度的变化; 分母部分表示图像边缘信息。初始深度和场景中边缘信息融合可以增加场景中边缘处深度估计效果, 使边缘更明显, 深度结果更准确。

3 实验结果

实验设备有: 北京卓立汉光微米级三维平移台 2 个, 维视工业相机 MV-VEM200SC 1 个。实验数据处理使用 Intel(R) Core i7-3770 计算机, 参数为 CPU3.40 GHz、内存 16.0 GB、Windows7 64 位, 处理软件为 MATLAB2016b 和 vs2010。

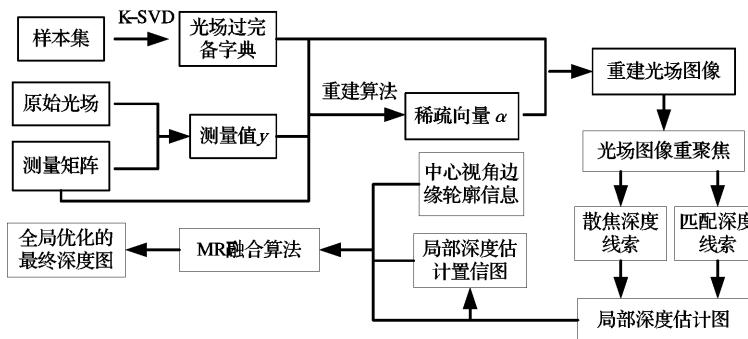


图 4 实验流程图

Fig. 4 Experimental flow chart

3.1 压缩光场仿真重建

利用三维平移台采集 5×5 视角真实光场数据组成原始光场数据, 每个光场子图像大小为 808×808 像素, 如图 5 所示。然后用 MATLAB 生成 808×808 像素周期排列的测量矩阵。将采集到的

光场子图像与相应的测量矩阵点乘再相加, 得到仿真的压缩传感测量值 y 。然后利用上述的重建算法重建出 5×5 视角的光场图像, 公式(5)中正则化参数取 2, 仿真重建结果如图 6 所示, 局部放大图像为 90×90 像素。

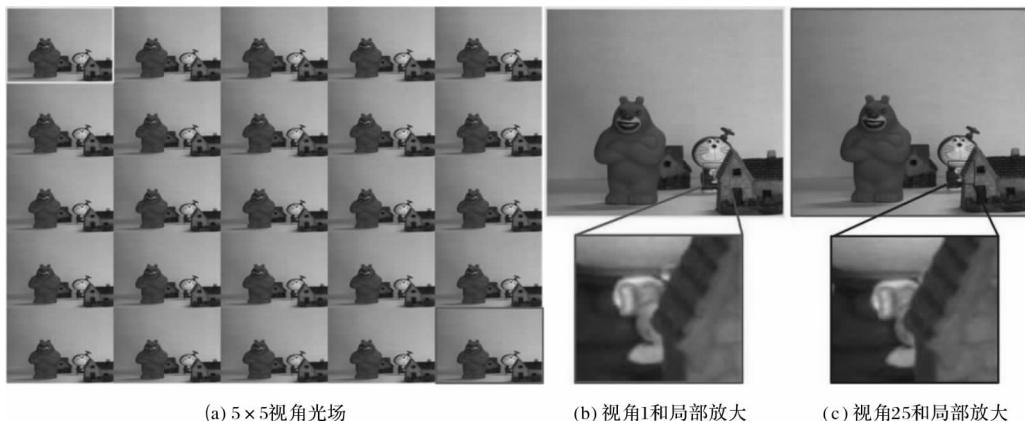


图 5 原始光场图

Fig. 5 Origin light field images

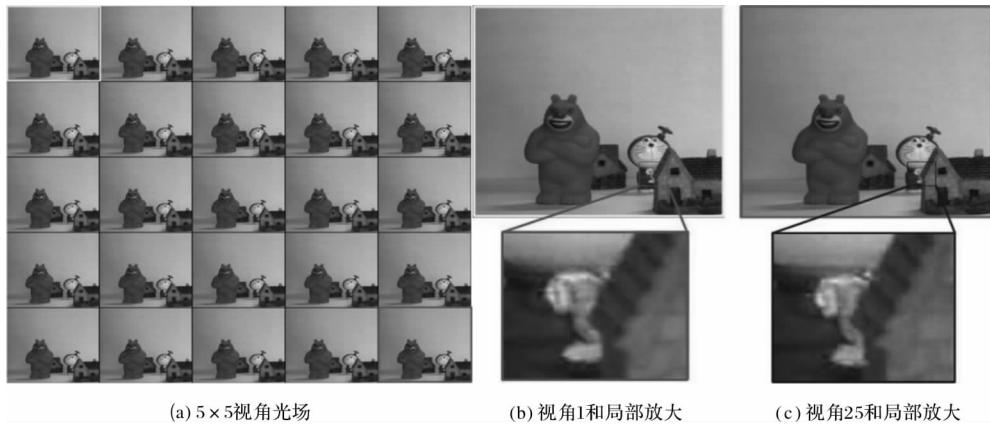


图 6 仿真重建光场图像

Fig. 6 Reconstruction light field images

通过图 5 和图 6 可以看出,仿真重建出的光场具有视差。但由于平移台精度以及实验操作时的精度不能达到亚像素级,所以造成重建光场清晰度有所下降,但是平滑效果好。

3.2 公开光场数据深度估计结果

为验证本文深度估计算法的准确性,用 HCI 大学光场数据“monasRoom”进行深度估计,分别用本文方法、Wang 方法^[19]、Tao 方法^[24]和 Mousnier 方法^[25]进行深度估计,结果如图 7 所示。从深度估计结果可以看出,与其他方法相比,本文方法的深度估计结果更好。在最终深度计算时,融合了图像的边缘信息,所以在场景边缘处深度估计结果更准确。

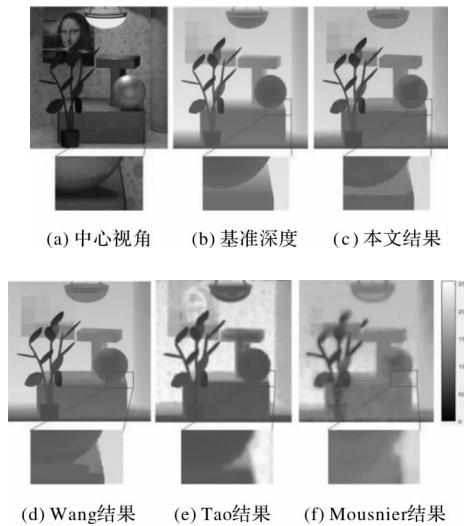


图 7 “monasRoom”深度估计结果

Fig. 7 Results of “monasRoom” depth estimation

在文献[19]的基础上,提出了融合初始深度、

置信度、边缘信息的深度估计算法,用边缘信息代替文献[19]深度正则化中遮挡预测的计算,在不影响计算精度的前提下,降低了时间开销。分别用本文和文献[19]对 HCI 大学的光场数据以及合成光场数据进行深度估计,其深度估计的结果与基准深度图做对比,用均方误差(mean squared error, MSE)作为误差评价标准,实验结果如表 1 所示。MSE 定义为

$$MSE = \frac{\sum \sum (d(x, y) - G(x, y))^2}{\sum \sum} \quad (13)$$

表 1 时间开销与 MSE 对比

Table 1 Comparison of time cost and MSE

光场数据	计算遮挡时间 ^[20] /s	本文计算轮廓时间/s	均方误差 ^[20]	本文均方误差
monosroom	46	<1	0.003 38	0.002 65
medieval	81	<1	0.012 72	0.012 94
papillion	43	<1	0.012 41	0.012 50
buddha	35	<1	0.009 56	0.009 88
outdoor	28	<1	0.002 76	0.002 86
sculptures	22	<1	0.006 02	0.004 45
plant	63	<1	0.012 43	0.012 41
livingroom	33	<1	0.004 39	0.004 41
bedroom	35	<1	0.012 10	0.009 60

从表 1 中时间复杂度对比中可看到:利用边缘信息替代遮挡线索优化初始深度图,降低了程序运行的时间开销,同时能够获取高精度的深度图。

3.3 原始光场和仿真重建光场深度估计结果

用本文的深度估计方法分别对采集的原始光

场、仿真重建出的光场进行深度估计,实验结果如图 8 所示。从图 8 可以看出,重建光场深度估计结果分层没有原始光场深度估计结果明显,原因在于重建光场的视差比原始光场有所减少。

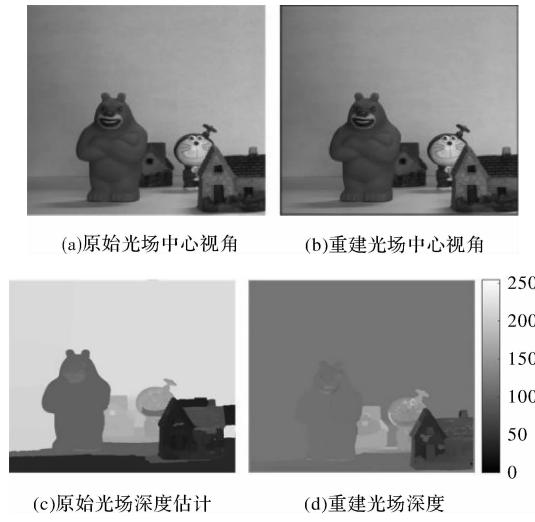


图 8 深度估计结果

Fig. 8 Results of depth estimation

为了验证重建光场深度估计准确度,用本文的深度估计方法分别对“monasRoom”和仿真重建的“monasRoom”进行深度估计,实验结果如图 9 所示。通过与基准深度对比,可以看出原始光场和仿真重建光场的准确性。

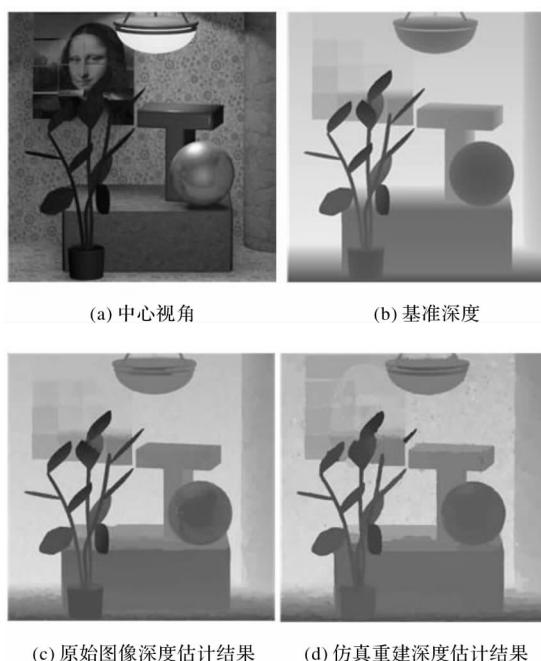


图 9 深度估计结果

Fig. 9 Results of depth estimation

4 结论

文中采用压缩感知原理仿真实现多视角光场重建,通过融合光场的初始深度、置信度、边缘信息实现光场深度估计,系统地完成了从仿真压缩光场采样获取光场数据,到重建光场,最后得到深度估计结果的整个过程。实验结果表明仿真重建光场可以用于深度估计。深度估计结果表明本文方法在场景边缘处有更好效果,边缘轮廓清晰,分层明显。实验验证了单个相机单次曝光便可重建出一组 5×5 视角的光场图像,将光场图像的采样率降低为 4%。为降低光场采集、处理和传输的数据量,提高边缘处深度估计结果的准确度提供了实验依据。为后续搭建更准确的物理平台,实现测量矩阵调制真实入射光线,提高重建光场的精度,为采集、重建、深度估计一体化奠定了实验基础,可用在三维重建、三维显示等领域。

参考文献:

- [1] LEVOY M, PAT H. Light field rendering [C]. US: ACM, 1996: 31-42.
- [2] NG R, LEVORY M, BREDIF M, et al. Light field photography with a hand-held plenoptic camera [J]. Computer Science Technical Report CSTR, 2005, 2 (11): 1-11.
- [3] YOON Y, JEON H G, YOO D, et al. Learning a deep convolutional network for light-field image super-resolution [C]. US: IEEE, 2015: 57-65.
- [4] SVEN W, BASTIAN GOLDLUECKE. Globally consistent depth labeling of 4D light fields [C]. US: IEEE, 2012: 41-48.
- [5] LI Nianyi, YE Jinwei, JI Yu, et al. Saliency detection on light field [C]. US: IEEE, 2014: 2806-2813.
- [6] SUN Shiyi, KUANG Cuifang, LIU Xu. Three dimensional surface reconstruction based on light field imaging [J]. Journal of Applied Optics, 2017, 38 (2): 210-214.
- [7] 孙试翼,匡翠方,刘旭. 基于光场成像的表面三维重构 [J]. 应用光学, 2017, 38(2): 210-214.
- [8] WILBURN B, JOSHI N, VAISH V, et al. High

- performance imaging using large camera arrays[J]. ACM Transactions on Graphics, 2005, 24 (3): 765-776.
- [8] MARWAH K, WETZSTEIN G, VEERARAGHAVAN A, et al. Compressive light field photography [J]. Acm Siggraph, 2012, 70(2): 49.
- [9] ASHOK A, NEIFELD M A. Compressive light-field imaging[J]. SPIE, 2010, 7690(1): 281-290.
- [10] BABACAN S D, ANSORGE R, LUSSI M, et al. Compressive light field sensing[J]. IEEE Transactions on Image Processing, 2012, 21 (12): 4746-4757.
- [11] CHANG L, YUAN Y, ZHANG X B, et al. Research on sampling and computer simulation of mask-based light field photography[J]. Acta Photonica Sinica, 2010, 39(6): 1104-1110.
常雷,袁艳,张修宝,等. 基于掩膜的光场成像仿真及采样研究[J]. 光子学报,2010,39(6):1104-1110.
- [12] LIU Y C, GONG H J, SHEN C L. Research of lightfield acquisition and reconstruction based on mask[J]. Acta Optica Sinica, 2014, 34(8): 105-110.
刘永春,龚华军,沈春林. 基于掩膜的光场采集与重建的研究[J]. 光学学报,2014,34(8):105-110.
- [13] QIN Yali, ZHANG Xiaoshuai, YU Linqian. Sparse sampling and reconstruction of compressive light field via low-rank matrix decomposition [J]. Opt. Precision Eng., 2017, 25(5): 1171-1177.
覃亚丽,张晓帅,余临倩. 基于低秩矩阵分解的光场稀疏采样及重构[J]. 光学精密工程, 2017, 25 (5): 1171-1177.
- [14] ZHANG Y, LV H, LIU Y, et al. Light-field depth estimation via epipolar plane image analysis and locally linear embedding [J]. IEEE Transactions on Circuits & Systems for Video Technology, 2017, 27 (4): 739-747.
- [15] CHEN C, LIN H, YU Z, et al. Light field stereo matching using bilateral statistics of surface cameras [C]. US: IEEE, 2014: 1518-1525.
- [16] HEBER S, POCK T. Shape from light field meets robust PCA[C]. GER: Springer International Publishing, 2014: 751-767.
- [17] JEON H G, PARK J, CHOE G, et al. Accurate depth map estimation from a lenslet light field camera[C]. US: IEEE, 2015: 1547-1555.
- [18] TAO M W, SRINIVASAN P P, HADAP S, et al. Shape estimation from shading, defocus, and correspondence using light-field angular coherence [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(3): 546-460.
- [19] 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.
- [20] DONOHO D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52 (4): 1289-1306.
- [21] LI Y, LI F, BAI B, et al. Image fusion via nonlocal sparse K-SVD dictionary learning[J]. Applied Optics, 2016, 55(7): 1814.
- [22] GISELSSON P, BOYD S. Diagonal scaling in douglas-rachford splitting and ADMM[C]. US: IEEE, 2014: 5033-5039.
- [23] YIN Xiaogen, ZHANG Xiaofang, ZHANG Weichao, et al. Study on 3D reconstruction based on light field digital refocusing[J]. Journal of Optoelectronics • Laser, 2015(5): 986-991.
尹晓良,张晓芳,张伟超,等. 基于光场数字重聚焦的三维重建方法研究[J]. 光电子·激光, 2015(5): 986-991.
- [24] TAO M W, HADAP S, MALIK J, et al. Depth from combining defocus and correspondence using light-field cameras[C]. US: IEEE Computer Society, 2013: 673-680.
- [25] MOUSNIER A, VURAL E, GUILLEMOT C. Partial light field tomographic reconstruction from a fixed-camera focal stack [J]. Computer Science, 2015, 1503: 01903.