

室内场景下弱纹理物体三维重建算法的研究

张庆鹏, 曹宇*

哈尔滨理工大学自动化学院, 黑龙江 哈尔滨 150080

摘要 针对室内小场景下图像采集位置受限、弱纹理物体重建效果不佳的问题, 提出了一种只需用手机采集图像的三维重建算法。首先, 用一种主动选择式图像匹配策略减少原始运动恢复结构算法中图像两两匹配的次数。然后, 将尺度不变特征变换(SIFT)算法改进为 Harris-SIFT 算法, 以提升算法的实时性; 通过全卷积神经网络获得预测深度并与多视图立体匹配算法进行融合, 以获得更多的稠密点云。最后, 用泊松表面重建算法完成物体的重建。实验结果表明, 本算法不仅能有效恢复室内场景下的物体细节特征, 对弱纹理物体表面的重建效果也较好。相比原始算法, 本算法所用的时间减少了 21.07%。

关键词 图像处理; 三维重建; 图像匹配; 运动恢复结构; 深度预测

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.0810017

Research on Three-Dimensional Reconstruction Algorithm of Weak Textured Objects in Indoor Scenes

Zhang Qingpeng, Cao Yu*

School of Automation, Harbin University of Science and Technology, Harbin, Heilongjiang 150080, China

Abstract In this paper, aiming at the problem of limited image acquisition location and poor reconstruction of weak texture objects in small indoor scenes, a three-dimensional reconstruction algorithm that only needs a mobile phone to acquire images is proposed. First, an active selective image matching policy is employed to reduce the number of images of pairwise matching in the original structure from motion algorithm. Then, the scale-invariant feature transform (SIFT) algorithm is improved to the Harris-SIFT algorithm to enhance real-time performance of the algorithm. Next, the predicted depth is obtained from the full consolidation neural network and fused with a multi-view stereo match algorithm to obtain more dense clouds. Finally, the reconstruction of the object is completed with a Poisson surface reconstruction algorithm. The experiment results show that the algorithm can not only effectively restore the detailed features of the object under the indoor scenes, but also has a better reconstruction effect on the surface of the weak texture objects. Compared with the original reconstruction algorithm, the time used by the algorithm is reduced by 21.07%.

Key words image processing; three-dimensional reconstruction; image matching; structure from motion; depth prediction

OCIS codes 100.6890; 150.1135; 110.2970

1 引言

某些特定室内场景空间狭小, 难以固定放置专业图像采集设备, 且存在采集设备移动困难、相机需

要预先标定等缺点, 但特定室内场景的三维重建却具有重要意义, 如博物馆文物的三维数字化^[1]、室内三维场景的在线展示^[2]。通过手机采集图像序列并以 5G 网络或 WiFi 的形式发送给上位机处理

收稿日期: 2020-07-30; 修回日期: 2020-08-11; 录用日期: 2020-09-15

基金项目: 黑龙江省普通本科高校青年创新人才培养计划(UNPYSCT-2015045)

* E-mail: cyhit@163.com

是一种新型三维重建方案,具有成本低、拍摄灵活和重建周期短等特点,目前有基于光流、基于机器学习和基于特征检测的三维重建算法。其中,基于特征检测的重建算法应用最广泛,但对弱纹理物体表面的重建精度不高^[3]。

人们对三维重建图像中的弱纹理问题进行了大量研究,张一飞等^[4]针对立体匹配中弱纹理区域匹配不准确的问题,提出了一种融合图像分割的匹配算法。徐雪松等^[5]提出了一种利用卷积神经网络(CNN)解决弱纹理或高光区域的立体匹配算法。刘一凡等^[6]提出了一种基于迭代最近点(ICP)与运动恢复结构(SFM)的双目立体视觉三维重建算法,最大限度地克服了两种算法的不足。林义闯等^[7]针对机器人视觉自主导航和避障在弱纹理场景中的问题,提出了基于激光投点器的解决方案。逢晖^[8]提出了一种基于 Harris 角点检测的弱纹理三维图像检测算法,可提高图像处理的质量。陈加等^[9]详细介绍了深度学习在单张图像物体三维重建中的应用及传统方法的研究现状。王芳等^[10]提出了一种能在缺帧环境下使用的弱纹理图像三维重建算法,可提升算法的图像纹理识别和修复重建能力。

针对传统三维重建算法在室内场景中对拍摄设备要求高、对弱纹理物体三维重建效率低且工作量大和重建时间长的问题,本文提出了一种基于特征提取与 CNN 的室内场景三维重建算法。将特定室内小场景作为研究对象,通过手机在不同角度拍摄 60 张图像,实现对室内场景物体的三维重建。

2 特征匹配算法

2.1 图像预处理

由于图像特征匹配算法无法在彩色图像上进行匹配,因此,需对彩色图像进行灰度化处理。室内场景中物体的颜色信息比较丰富,灰度化必然会损失

部分彩色信息,如直接灰度化算法会损失一部分图像原有的彩色信息,直方图均衡化算法通过改变图像灰度分布使图像灰度级分布更均匀,从而增强图像的细节信息,但会改变图像亮度。为了更好地保留图像的彩色信息,采用一种融合光照和彩色信息的灰度图像补偿算法对图像进行灰度化^[11]。

2.2 改进的 Harris-尺度不变特征变换匹配算法

Harris-尺度不变特征变换(SIFT)算法^[12]的实时性较好,但准确率一般。为了提高算法的准确率,首先用熵值法对提取的特征角点进行提纯;然后将欧氏距离改为马氏距离,用随机抽取一致性算法替换筛选环节的阈值法;最后通过实验验证了本算法在室内场景下的图像匹配准确率。

1) 特征角点提取

首先用 Harris 算法从预处理的灰度图像中提取特征角点,可表示为

$$M(x, y) = \text{Det}[\mathbf{H}(x, y)] - k \cdot \{\text{Tr}[\mathbf{H}(x, y)]\}^2, \quad (1)$$

式中, (x, y) 为像素点坐标, $M(x, y)$ 为角点响应函数, \mathbf{H} 为 Hessian 矩阵, $\text{Tr}(\mathbf{H})$ 为矩阵 \mathbf{H} 的迹, $\text{Det}(\mathbf{H})$ 为矩阵 \mathbf{H} 的行列式值, k 为常量,实验取 $k=0.04$,角点响应函数的局部极大值点即为角点。一般来说,图像的熵与图像包含的信息大小成正比例关系。因此,用熵值法^[13]对原 Harris-SIFT 算法提取出的特征角点进行提纯,即删除熵值小的特征角点,从而提高算法的匹配准确率,同时减少匹配所用的时间,提高算法的实时性。

2) 生成 SIFT 特征描述子

Harris 算法能获得特征角点的位置信息,精确定位所有的特征角点。但 SIFT 算法对图像进行匹配时,还需获得特征角点的方向信息。求取特征角点的方向信息时,首先要计算出图像像素点的所有梯度方向和模值,可表示为

$$m(x, y) = \{ [L(x+1, y) - L(x, y-1)]^2 + [L(x, y+1) - L(x, y-1)]^2 \}^{\frac{1}{2}}, \quad (2)$$

$$\theta(x, y) = \arctan\{ [L(x, y+1) - L(x, y-1)] / [L(x+1, y) - L(x-1, y)] \}, \quad (3)$$

式中, $m(x, y)$ 为 (x, y) 处梯度的模值, $\theta(x, y)$ 为 (x, y) 处梯度的方向, $L(x, y)$ 为所有尺度的关键点。根据 3σ (σ 为尺度空间因子) 采样原则,梯度模值用 1.5σ 的高斯分布加权,将邻域窗口半径设置为 $3 \times 1.5\sigma$,并通过梯度直方图统计邻域范围内角点的梯度和方向,从而生成所需的 128 维特征描述子。

3) 特征匹配

用马氏距离替代原始 Harris-SIFT 算法的欧氏距离,以判定两个特征向量之间的相似性。马氏距离是一种数据的协方差距离,可计算两个未知样本集的相似性,相比欧氏距离,考虑了不同特性之间的联系,且与尺度无关,即独立于测量尺度。两个向量之间的马氏距离可表示为

$$D_{xy} = \sqrt{(X - Y)^T S^{-1} (X - Y)}, \quad (4)$$

$$S = E \{ [X - E(X)] [Y - E(Y)] \}, \quad (5)$$

式中, X 、 Y 为两组向量, S 为两组向量的协方差矩阵, S^{-1} 为协方差逆矩阵, E 为均值。可以发现, 相比欧氏距离, 马氏距离多了一个协方差逆矩阵, 能更准确地衡量距离关系, 当协方差矩阵为单位矩阵时则演化为欧氏距离。虽然增加了运算量, 但马氏距离排除了量纲的影响, 具有尺度无关性; 同时也排除了特征向量之间相关性的干扰, 匹配精度更高, 整体效果优于欧氏距离。为进一步提高算法的匹配准确率, 需剔除错误的匹配, 因此, 将阈值法改进为随机抽样一致性(RANSAC)^[14]算法, 可在含有很多外部点的数据中精确估计模型参数。

2.3 室内小场景匹配实验

实验环境: 系统为 Windows10, 软件为 Matlab2018b, 三维模型通过 Meshlab2016 展示。实验在同一台 PC 机下进行, PC 机的配置: 处理器为英特尔 i7-6700HQ, 主频为 2.6 GHz, 内存为 8 G, 硬盘为 500 G 固态硬盘, 显卡为英伟达 GTX960M 型号的 4G 独立显卡。拍摄设备为 iPhone6 手机, 随机从室内场景拍摄的 60 张图像中选取 2 张图像进行匹配实验, 待匹配图像如图 1 所示。

匹配准确率 P 可表示为

$$P = 1 - \frac{N_F}{N_T}, \quad (6)$$

式中, N_F 为错误匹配点对, N_T 为所有的匹配点对。表 1 为用图像灰度化补偿算法处理后 4 种匹配算法的实验数据, 4 种匹配算法的实验结果如图 2 所示。可以发现, 相比其他算法, 本算法的性能指标更优。



图 1 手机拍摄的待匹配图像

Fig. 1 Image to be matched taken by mobile phone

原因是本算法用熵值法筛选 Harris-SIFT 算法提取的特征角点, 减少了后续匹配环节的工作量, 并用马氏距离代替欧氏距离。相比原始 Harris-SIFT 算法, 匹配准确率提升了 1.25 个百分点, 所用时间减少了 8.38%。

表 1 4 种匹配算法的准确率和运行时间

Table 1 Accuracies and running time of 4 matching algorithms

Matching algorithm	Harris	SIFT	Harris-SIFT	Improved Harris-SIFT
Left image feature point	2057	2786	1490	1450
Right image feature point	2172	2801	1612	1597
Matching point	1600	821	506	492
Matching accuracy / %	81.61	93.54	93.67	94.92
Matching time / s	5.02	11.64	3.94	3.61

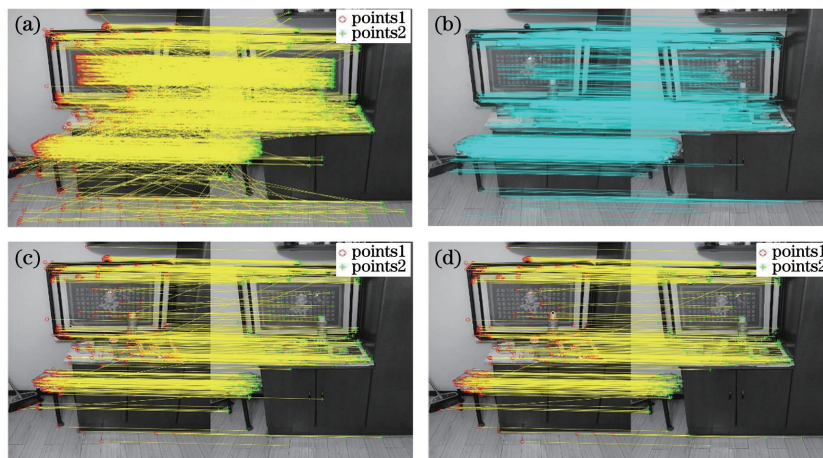


图 2 不同算法的匹配结果。(a) Harris 算法; (b) SIFT 算法; (c) Harris-SIFT 算法; (d) 本算法

Fig. 2 Matching results of different algorithms. (a) Harris algorithm; (b) SIFT algorithm; (c) Harris-SIFT algorithm;

(d) our algorithm

3 室内特定小场景物体的三维重建

3.1 稀疏点云重建

将手机拍摄的 60 张图像序列用 SFM 算法对特定场景进行三维重建。首先,用本算法匹配,将包含最多匹配点的图像作为初始图像对,通过计算获得相机的运动信息和场景中物体的结构信息。其次,用三角测量法得到场景的初始稀疏点云,并不断增加新的图像进行迭代。最后,用光束平差(BA)法优化参数,获得稀疏的空间三维点云模型。

3.2 主动选择式图像匹配策略

由于 SFM 算法需要对所有图像进行两两匹配,实时性较差,为了提高图像的匹配效率,提出了一种用于室内场景的主动选择式图像匹配策略。

SFM 算法选择初始的 2 张匹配图像时,会选择匹配特征最多的图像,即拍摄图像位置越近,获得的特征越多。假设拍摄了 n 张图像,则匹配时选取一定数量的图像,获得足够数量的匹配点进行三维重建,选取的最邻近图像数量为

$$p = \frac{n}{l}, \quad (7)$$

式中, l 为正整数。不同 l 时的实验结果如图 3 所示,具体参数如表 2 所示。可以发现, l 越大,匹配次数越少,所用时间也越少。当 $l=2$ 时,重建后的点云数量没有明显下降;当 $l=3$ 时,虽然消耗的时间更少,但点云数量有明显下降,稀疏点云的恢复并不理想。因此最终选取 $l=2$,即 SFM 算法的匹配环节中每张图像使用最邻近的 30 张图像进行匹配。

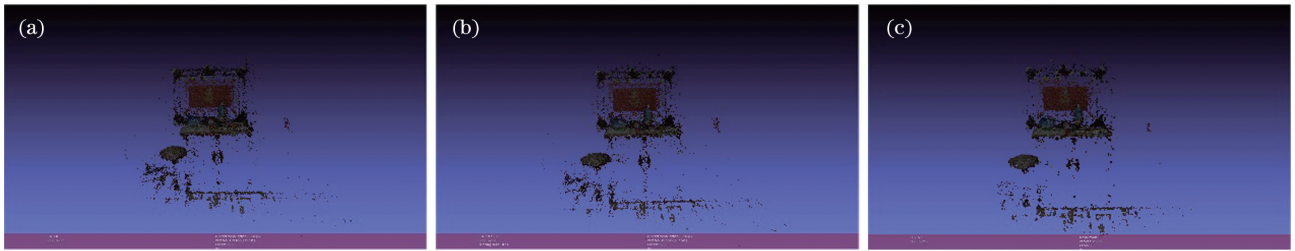


图 3 不同 l 时的实验图像。(a) $l=3$;(b) $l=2$;(c) $l=1$

Fig. 3 Experimental images at different l . (a) $l=3$; (b) $l=2$; (c) $l=1$

表 2 不同 l 时的图像参数

Table 2 Image parameters at different l

l	Number of image matches	Match time /s	Number of point clouds
1	32719	94	101823
2	24589	61	99745
3	15541	42	65521

3.3 稠密点云重建

稠密重建(MVS)是一种将 SFM 稀疏点云稠密化的算法,利用 SFM 算法输出估计图像序列中像素点的深度值,然后融合深度图获得稠密三维点云。MVS 算法将传统立体匹配的两视图扩展到多视图,但对于弱纹理物体表面的三维重建,深度估计误差

较大,原因是表面的灰度信息有限。因此,利用基于残差网络(ResNet-50)的全卷积神经网络算法预测图像,通过 MVS 算法设定阈值,将预测的深度融合到 MVS 深度图中,解决了弱纹理目标深度图误差较大的问题。

为保证模型收敛,用文献[15]提出的全卷积神经网络,即 ResNet50 深度网络模型预测图像深度。ResNet50 在构建深层网络时减弱了退化现象,因此在反卷积层输出后增加了修正线性单元(ReLU),可表示为

$$X_{\text{ReLU}}(a) = \begin{cases} a, & a > 0 \\ 0, & a \leq 0 \end{cases} \quad (8)$$

图 4 为用于预测深度的全卷积网络结构,为了保

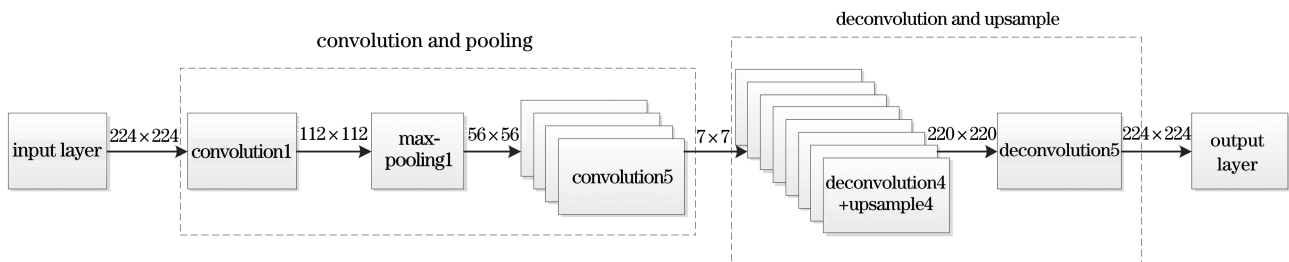


图 4 ResNet50 的结构

Fig. 4 Structure of the ResNet50

证系统的实时性,去掉 ResNet50 中的平均池化层和全连接层,仅保留卷积层,以提取图像中包含的特征信息,最终获得尺寸为 $7 \times 7, 2048$ 通道的特征图。

用 3 种常见的预测评价指标验证全卷积神经网络模型预测深度的有效性,包括均方根误差 (X_{RMSE})、平均相对误差 (Y_{REL}) 和平均 \log_{10} 误差 ($Z_{\log_{10}}$),可表示为

$$X_{RMSE} = \sqrt{\frac{\sum_{i=1}^N (y_i - d_i)^2}{N}}, \quad (9)$$

$$Y_{REL} = \frac{\sum_{i=1}^N \frac{|y_i - d_i|}{y_i}}{N}, \quad (10)$$

$$Z_{\log_{10}} = \frac{\sum_{i=1}^N |\log_{10} y_i - \log_{10} d_i|}{N}, \quad (11)$$

式中, y_i 为第 i 个像素点的真实深度, d_i 为第 i 个像素点深度的预测值, N 为物体像素点总数。表 3

表 3 不同算法的误差

Table 3 Errors of different algorithms

Algorithm	X_{RMSE}	Y_{REL}	$Z_{\log_{10}}$
MVS	2.216	0.101	0.075
Ours	1.123	0.081	0.038

为本算法与 MVS 算法的误差,可以发现,相比 MVS 算法,本算法在包含弱纹理物体的三维场景重建中,3 种误差评价指标均更小。

4 实验结果与分析

实验共获得 60 张不同角度的图像,部分图像如图 5 所示,通过 60 张图像获得的稀疏点云如图 6 所示。



图 5 部分实验图像

Fig. 5 Some experimental images

通过 SFM 算法获得稀疏点云和图像的三维空间位置后,用 CNN 融合 MVS 算法,以减少弱纹理物体表征效果差的问题,使场景的点云更稠密,重建效果如图 7 所示,用泊松表面重建算法得到的重建图像如图 8 所示。

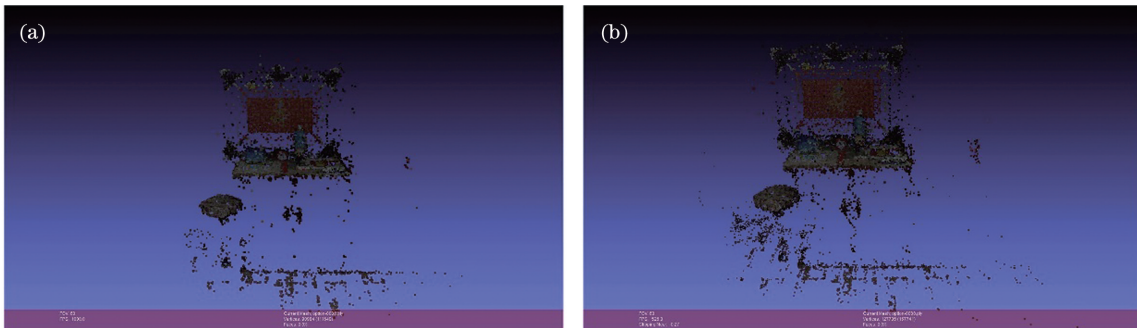


图 6 稀疏三维点云。(a) SFM 算法;(b)改进的 SFM 算法

Fig. 6 Sparse three-dimensional point cloud. (a) SFM algorithm; (b) improved SFM algorithm

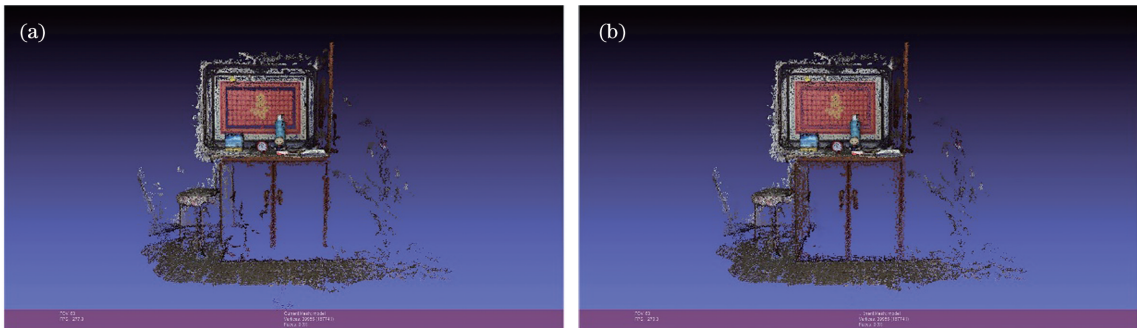


图 7 物体的稠密点云。(a) MVS 算法;(b)本算法

Fig. 7 Dense point cloud of the object. (a) MVS algorithm; (b) our algorithm

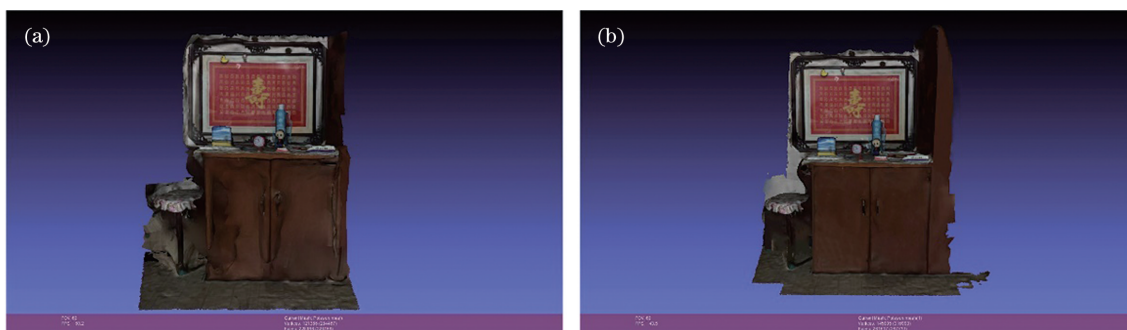


图 8 泊松表面的三维重建结果。(a)原始算法;(b)改进的算法

Fig. 8 Three-dimensional reconstruction results of the Poisson surface. (a) Original algorithm; (b) improved algorithm

本算法由于提取了更多的稠密点云,泊松表面重建算法最终恢复的场景细节效果更优。2种算法的运行时间、面片数量和结构相似度(SSIM)如表4所示。可以发现,相比原始算法,本算法所用的时间减少了21.07%,获得的点云数量增加了14.66%,场景SSIM提升了52.83%。

表4 不同算法的三维重建结果

Table 4 Three-dimensional reconstruction results of different algorithms

Algorithm	Time /min	Point cloud	SSIM
SFM+MVS	8.21	228166	0.53
Improve algorithm	6.48	261617	0.81

5 结 论

针对特定室内小场景提出了一种无标定三维重建算法,首先,用手机采集图像序列,解决了拍摄空间受限的问题。然后,通过改进匹配算法和匹配策略,提升了SFM算法的实时性。其次,用全卷积神经网络结构以及多层卷积结构进行特征学习,获得与输入图像尺寸相同的深度图,通过设定阈值将其融入MVS深度图中,解决了弱纹理物体表面的重建难题。最后,用泊松表面重建算法对室内三维场景进行重建。实验结果表明,本算法在特定室内场景下无需标定,且只需使用手机采集图像,对图像序列的数目也没有严格要求,可解决室内小场景下弱纹理物体重建效果不佳的问题。

参 考 文 献

[1] Tu D W, Lan H, Zhang X. Digital three-dimensional reconstruction technology of cultural relics [J]. Laser & Optoelectronics Progress, 2019, 56(19): 191504.
屠大维, 兰浩, 张曦. 文物数字化三维重建技术研究[J]. 激光与光电子学进展, 2019, 56(19):

191504.

- [2] Zhang H X, Fang Y T, Li M. Robust reconstruction method of 3D room layout with visual-inertial module [J]. Journal of Computer-Aided Design & Computer Graphics, 2020, 32(2): 262-269.
张宏鑫, 方雨桐, 利明. 结合视觉惯性模组的室内三维布局鲁棒重建方法[J]. 计算机辅助设计与图形学学报, 2020, 32(2): 262-269.
- [3] Qiao Y J, Zhang S Y, Zhao Y H. Surface robust reconstruction method for high light and weak textured objects [J]. Acta Photonica Sinica, 2019, 48(12): 1212002.
乔玉晶, 张思远, 赵宇航. 高光弱纹理物体表面鲁棒重建方法[J]. 光子学报, 2019, 48(12): 1212002.
- [4] Zhang Y F, Li X F, Tian X D. Stereo matching algorithm based on image segmentation [J]. Journal of Computer Applications, 2020, 40(5): 1415-1420.
张一飞, 李新福, 田学东. 基于图像分割的立体匹配算法[J]. 计算机应用, 2020, 40(5): 1415-1420.
- [5] Xu X S, Wu J J. Stereo matching algorithm based on multiscale fusion [J]. Pattern Recognition and Artificial Intelligence, 2020, 33(2): 182-187.
徐雪松, 吴俊杰. 多尺度融合的立体匹配算法[J]. 模式识别与人工智能, 2020, 33(2): 182-187.
- [6] Liu Y F, Cai Z J. Binocular stereo vision three-dimensional reconstruction algorithm based on ICP and SFM [J]. Laser & Optoelectronics Progress, 2018, 55(9): 091503.
刘一凡, 蔡振江. 基于ICP与SFM的双目立体视觉三维重构算法[J]. 激光与光电子学进展, 2018, 55(9): 091503.
- [7] Ling Y M, Lü N G, Lou X P, et al. Robot vision system for 3D reconstruction in low texture environment [J]. Optics and Precision Engineering, 2015, 23(2): 540-549.
林义闽, 吕乃光, 娄小平, 等. 用于弱纹理场景三维重建的机器人视觉系统[J]. 光学精密工程, 2015, 23(2): 540-549.
- [8] Pang H. Weak texture three-dimensional

- discontinuous image detection based on Harris corner [J]. *Computer Simulation*, 2016, 33(9): 431-434.
逢晖. 基于 Harris 角点的弱纹理三维非连续图像检测 [J]. *计算机仿真*, 2016, 33(9): 431-434.
- [9] Chen J, Zhang Y Q, Song P, et al. Application of deep learning to 3D object reconstruction from a single image [J]. *Acta Automatica Sinica*, 2019, 45(4): 657-668.
陈加, 张玉麒, 宋鹏, 等. 深度学习在基于单幅图像的物体三维重建中的应用 [J]. *自动化学报*, 2019, 45(4): 657-668.
- [10] Wang F, Wang W. Three dimensional reconstruction method of weak texture image under the condition of lack of frame [J]. *Journal of Xi'an Polytechnic University*, 2016, 30(4): 477-482.
王芳, 汪伟. 缺帧环境下弱纹理图像的三维重建方法 [J]. *西安工程大学学报*, 2016, 30(4): 477-482.
- [11] Yang L, Hu L K, Wu H N. Scale invariant feature transform image matching algorithm combining color and exposure information [J]. *Journal of Guangxi University (Natural Science Edition)*, 2018, 43(6): 2197-2205.
阳兰, 胡立坤, 吴浩楠. 融合彩色和光照信息的尺度不变特征变换图像匹配算法 [J]. *广西大学学报(自然科学版)*, 2018, 43(6): 2197-2205.
- [12] Zhang Q P, Cao Y. Image matching algorithm based on exposure and color information [J]. *Laser & Optoelectronics Progress*, 2019, 56(19): 191004.
张庆鹏, 曹宇. 一种融合光照和彩色信息的图像匹配算法 [J]. *激光与光电子学进展*, 2019, 56(19): 191004.
- [13] Shang M S, Wang K C. Advanced image registration method based on Harris and SIFT algorithm [J]. *Microelectronics & Computer*, 2018, 35(6): 132-134, 140.
尚明珠, 王克朝. 一种改进的 Harris 与 SIFT 算子结合的图像配准算法 [J]. *微电子学与计算机*, 2018, 35(6): 132-134, 140.
- [14] Wang X, Liu Y L, Li L Y. Symmetry-based projective algorithm for 3D localization of small target at a wide field of view [J]. *Acta Optica Sinica*, 2020, 40(10): 1015001.
王潇, 刘育梁, 李丽艳. 基于对称性投影的大视场小目标三维定位算法 [J]. *光学学报*, 2020, 40(10): 1015001.
- [15] Tao Y Z. 3D reconstruction of texture-less object using incremental structure from motion [D]. Wuhan: Wuhan University, 2018: 31-45.
陶易之. 基于增量式运动恢复结构的弱纹理目标三维重建 [D]. 武汉: 武汉大学, 2018: 31-45.