

# 基于深度相机的古建筑木结构裂缝识别 联合多分辨率三维重建方法

阎卫东,石日新,王井利,刘国奇,马健

(沈阳建筑大学交通与测绘工程学院,辽宁 沈阳 110168)

**摘要** 目的 研究木结构古建筑的三维数字化技术以及木结构表面裂缝分析方法,为该类古建筑的裂缝测量和维护提供参考。**方法** 通过 Mask R-CNN 对 RGBD 图像中的裂缝信息进行识别,结合位姿优化策略计算每一帧的相机姿态,进而获得裂缝的三维信息;构造一种包含古建筑三维模型与裂缝信息的多分辨率三维重建模型,基于所构造的模型分析古建筑表面裂缝的长度和深度。**结果** 提高裂缝三维重建精度前,长度平均计算误差为 12.3%,深度平均计算误差为 36.8%;提高裂缝三维重建精度后,长度平均计算误差为 9.3%,深度平均计算误差为 19.7%。**结论** 提出的识别方法可以快速、精准地识别古建筑木结构中的裂缝并获得古建筑的三维模型,相较于传统裂缝测量效率更高、成本更低。

**关键词** 古建筑;木结构;裂缝识别;三维重建;多分辨率;深度相机;帧块

中图分类号 TU398.1 文献标志码 A

## Crack Recognition and Multi-resolution 3D Reconstruction of Ancient Wooden Structures Based on Depth Camera

YAN Weidong, SHI Rixin, WANG Jingli, LIU Guoqi, MA Jian

(School of Transportation and Geomatics Engineering, Shenyang Jianzhu University, Shenyang, China, 110168)

**Abstract:** This paper was proposed to study the 3D digital technology and surface crack analysis method of wooden structure ancient buildings, and to provide reference for the modeling and crack measurement of this kind of ancient buildings. On the basis of Mask R-CNN to identify the crack information in RGBD images, the local to global optimization strategy was used to calculate the camera pose, and a multi-resolution implicit structure to contain the 3D model of ancient architecture and crack information and analyze the length and depth of cracks on the surface of ancient buildings. A multi-resolution 3D model with fracture information was obtained, which improved the accuracy of 3D reconstruction of the fracture, and analyzed the length and depth of the fracture. In

收稿日期:2023-02-15

基金项目:国家自然科学基金项目(51978420);辽宁省教育厅科研项目(LJKZ0602, LNQN202035)

作者简介:阎卫东(1964—),男,教授,博士,主要从事测绘、建筑工程安全管理等方面研究。

Red Wood data set and real scenes, the average calculation error of fracture length and depth is 12.3% and 36.8% respectively before improving the 3D reconstruction accuracy. After improving the 3D reconstruction accuracy, the average calculation error of length and depth is 9.3% and 19.7% respectively. It is proved that the method can quickly and accurately identify cracks in ancient wooden structures and obtain 3D models of ancient buildings, which is more efficient and cheaper than traditional crack measurement.

**Key words:** ancient architecture; wooden structure; crack recognition; 3D reconstruction; multi-resolution; RGB-D camera; frame patch

由于年久失修、气候干扰、人为损伤等原因,古建筑中的木结构会发生各种形式的破坏,其中裂缝是最常见的损伤形式<sup>[1]</sup>。且随着时间推移会出现不可逆转的结构损坏,甚至造成建筑物坍塌,因此对古建筑木结构裂缝检测具有重要意义。

传统人工测量存在工作难度大、效率低等问题,因此学者们提出了一系列基于图像识别深度学习的方法来检测图像中的裂缝,其中以区域分割方法最为经典<sup>[2-3]</sup>。R. Girshick<sup>[4]</sup>在Faster R-CNN网络中添加预测物体掩码信息的分支,K. He等<sup>[5]</sup>提出了经典的Mask R-CNN实例分割网络,实现了图像的目标检测。马健等<sup>[6]</sup>使用YOLOv5来识别图像中的裂缝。与YOLO等目标检测网络相比,Mask R-CNN包含的“实例”信息可以更精确地刻画裂缝的区域,而不仅仅是裂缝处的包围盒。韩士伟等<sup>[7-8]</sup>利用三维重建技术将古建筑的整体结构、表面纹理等信息进行重建,获得裂缝检测工作中所需要的裂缝大小、深度等具体信息。韦虎等<sup>[9-10]</sup>利用双目立体视觉与多视几何技术恢复古建筑表面的三维结构与纹理,精度提高,但计算速度慢。阮競芸等<sup>[11-13]</sup>使用深度相机拍摄多视角深度图,利用立体匹配与特征匹配技术计算各视角下的相机位姿,再将深度图融合到三维空间,其特点是速度快,可以实时地完成重建,更适合大规模古建筑的三维重建工作,但其精度较低。S. Izadi等<sup>[14]</sup>使用栅格化的方式表示整个三维场景,利用高速的GPU设备,通过截断距离函数计算空间中每一个栅格到其最近表面的距离。A. Dai

等<sup>[15]</sup>在M. Nießner等<sup>[16]</sup>提出的Voxel Hashing算法基础上改进,提出了Bundle Fusion三维重建系统,引入从局部到全局的相机位姿优化策略,在重建大型建筑上具有更好的鲁棒性,但此类三维重建方法,仅生成一个由点云构造的三维模型,不能识别分析模型中包含的裂缝信息。J. McCormac等<sup>[17-19]</sup>提出将图像识别和三维重建联合的方法,利用CNN网络识别场景中不同语义的区域,可将识别到的物体融合到三维模型中。这类方法重建精度较低,很难将裂缝刻画清晰。

综合以上研究,笔者使用Mask R-CNN实例分割网络提取古建筑木结构表面的裂缝信息,使用基于Bundle Fusion系统的三维重建框架将裂缝信息与图像信息应用到三维重建中,建立一个含有裂缝信息的古建筑三维结构模型。进而提出一个裂缝分析方法,在生成含裂缝信息的三维模型后,计算裂缝的长度和深度。将Voxel Hashing算法中体素的结构推广到不同分辨率的应用场景中,以实现速度快、精度高的三维重建。

1 裂缝识别与多分辨率三维重建方法

图1描绘了裂缝识别与多分辨率三维重

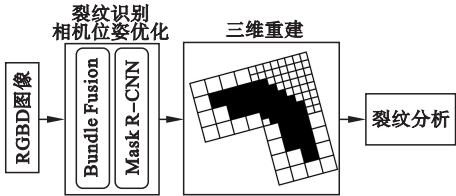


图1 系统框架

Fig. 1 The framework of system

建方法系统框架,包括裂缝识别与相机位姿优化模块、三维重建模块和裂缝分析模块。

1.1 裂缝识别与相机位姿优化模块

图2为裂缝识别与相机位姿优化模块,该模块的输入为一个RGBD视频流,裂缝识别与相机位姿优化在此模块的两个线程中并行执行。

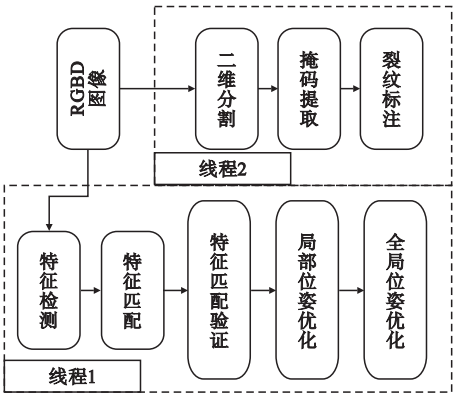


图2 裂缝识别与相机位姿优化模块  
Fig. 2 The crack recognition and camera pose optimization module

1.1.1 相机位姿优化

相机位姿是每一帧RGBD图像的相机拍摄位置,是三维重建技术的基本要素,一般通过优化的方法计算得到,相机位姿优化在一个线程中循环执行。当一个新的RGBD帧 $I_t(t \in [0, n], n$ 为当前相机拍摄的总帧数)到达时,利用SIFT<sup>[20]</sup>特征提取方法检测 $I_t$ 图像的特征点集合 $\delta_t$ ,并寻找 $\delta_t$ 与 $\delta_{t-1}$ 的特征匹配关系,使用Bundle Fusion中的特征点验证方法来确保特征点匹配关系的可靠性。通过特征点对应关系,利用对极几何约束计算 $I_t$ 和 $I_{t-1}$ 之间的初始相机位姿 $T'_{init}$ , $T'_{init}$ 仅用于 $I_t$ 帧的首次融合, $I_t$ 帧的重新融合过程采用优化后的相机位姿。当 $t$ 为10的整数倍时,将最后到达的10帧构造为一个帧块,在帧块内进行局部相机位姿优化,在帧块之间进行全局相机位姿优化。优化过程中,构造特征点光度项与几何项的重投影误差使用高斯牛顿法求解。优化的结果为每一帧相机的位姿,记为 $T'_{otm}$ 。

1.1.2 裂缝识别

每间隔 $k$ 帧进行一次裂缝识别,以使相机位姿优化与裂缝识别速度一致。对于到达的RGBD帧 $I_t(t \bmod k = 0)$ ,将其中的RGB图像输入到Mask R-CNN网络中预测其实例信息 $M_t$ 。

$M_t$ 和 $I_t$ 尺度一致,记录了实例所属的类别以及相应的位置,即 $M_t(x, y) = m$ ,其中 $m \in \{0, c\}$ 为像素的掩码值, $c$ 为Mask R-CNN网络训练时裂缝的标签。 $M_t$ 中存储了 $I_t$ 每一个像素对应的裂缝信息,此信息与 $I_t$ 中颜色和深度信息都要被融合到TSDF模型中。由于裂缝识别模块与相机位姿优化模块在不同的线程中执行且裂缝识别并不是在每一帧都执行,因此当 $I_t$ 进行融合时, $M_t$ 可能还未计算得到。为了解决这一问题,将 $M_t$ 信息存储在程序中,在重新融合 $I_t$ 时判断其是否有对应的 $M_t$ ,若存在则融合 $M_t$ 。

1.2 三维重建模块

1.2.1 三维重建模型结构

在Voxel Hashing算法的基础上,新增一个多分辨率模型来存储重建的三维结构信息和裂缝信息。此模型使用体素块来表达三维场景,体素块分为基本块和细化块,基本块中存储初始重建的3D数据和裂缝的实例信息,细化块中存储细化的3D数据。图3展示多分辨率重建模型的结构。使用哈希表和细化表来管理基本块和细化块。每一个基本块和细化块都对应真实场景的一个三维坐

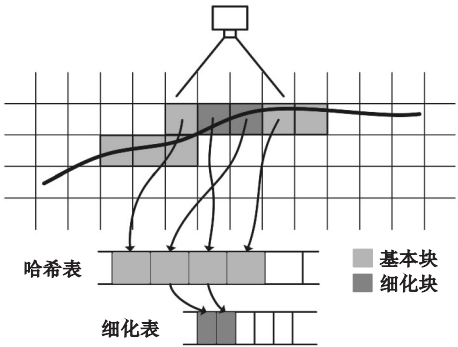


图3 多分辨率重建模型

Fig. 3 The multi-scale reconstruction module

标。每个基本块和细化块的边长为 0.08 m, 在当前的实现中, 一个基本块由  $8^3$  个体素组成, 一个细化块由  $32^3$  个体素组成, 基本块中相邻体素之间的距离为 10 mm, 细化块中相邻体素之间的距离为 2.5 mm。

当拍摄到一帧后, 根据此帧的相机位姿计算位于  $I_t$  相机视野下的基本块, 并获得这些基本块中体素的三维坐标  $W = (x, y, z)$ , 通过哈希函数计算哈希条目:

$$H = (p_1x + p_2y + p_3z) \bmod n. \quad (1)$$

式中:  $p_1, p_2, p_3$  为最大的素数;  $n$  为哈希表的大小。每一个基本块的内存位置可以由  $H$  索引, 每一个体素的内存位置可以由与其所属基本块的坐标偏差计算得到。将基本块中的体素投影到  $I_t$  的成像平面, 读取  $I_t$  中 RGB 图像、深度图像和掩码图像投影点的值, 计算投影点处相应的信息, 即 SDF、权重、颜色和掩码信息。若投影点恰好是裂缝所在的位置, 则细化该基本块——即在细化表中顺序查找一个空闲的细化条目, 将此细化条目存储在  $I_t$  相机视野下的基本块中, 再将细化块中的体素投影到  $I_t$  的成像平面, 计算投影点处对应的 SDF、权重、颜色和掩码信息。

### 1.2.2 含语义信息的体素融合

当完成  $I_t$  帧的裂缝识别与相机位姿优化后, 计算位于  $T_{\text{init}}'$  视野下的基本块体素的世界坐标。通过式(1)建立世界坐标和基本块体素之间的关联, 并将 RGB 图像和深度图像中的信息融合到基本块中的体素中。每个体素都包含 SDF 值  $D$ 、权重值  $E$ 、颜色值  $C$  和掩码值  $M$ 。

将基本块中每个体素的世界坐标  $W$  投影到  $I_t$  的成像平面上, 得到二维投影点  $p$ :

$$p = (T_{\text{opt}}'^{-1} \times W)_{3 \times 3} \times T_c. \quad (2)$$

式中:  $T_c \in \mathbb{R}^{3 \times 3}$  为相机的内参矩阵。新计算的 SDF 值  $D'$ 、权重值  $E'$ 、颜色值  $C'$  见式(3)~(5):

$$D' = \text{Depth}(p) - W_z. \quad (3)$$

$$E' = 1 - \text{Depth}(p)/d_{\text{max}}. \quad (4)$$

$$C' = \text{RGB}(p). \quad (5)$$

式中:  $d_{\text{max}}$  为融合允许的最大深度值;  $W_z$  为  $W$  在  $z$  轴上的距离;  $\text{Depth}()$  函数和  $\text{RGB}()$  函数为获取  $I_t$  对应的深度图和 RGB 图中指定像素的深度值和颜色值。然后将  $D'$ 、 $E'$ 、 $C'$  的值加权融合到  $D$ 、 $E$ 、 $C$  中。

重新融合在全局位姿优化模块中相机位姿变化最大的 10 帧, 并判断这些帧是否有对应的掩码图。更新操作与 Bundle Fusion 中一致, 首先根据  $T_{\text{init}}'$  删除之前融合的信息, 再根据  $T_{\text{opt}}'$  将新算得的信息融合到细化块内的体素中, 获取带有掩码信息的帧, 将基本块中体素投影到掩码图中获得掩码值:

$$M' = \text{Mask}(x, y). \quad (6)$$

$\text{Mask}()$  函数为获取掩码图中指定像素的掩码值, 将  $M'$  融合到  $M$  中, 得出:

$$M = \begin{cases} M', & m > 0, D \leq 0.01; \\ M, & \text{其他}. \end{cases} \quad (7)$$

细化  $M > 0$  的基本块, 即通过 1.2.1 节中的方法, 建立细化块中体素之间的关联和存储位置, 并按照上述方法更新细化块中体素的值。最终得到一个包含裂缝信息的木结构三维模型, 模型中以体素的形式表达了木结构的形状、纹理与裂缝信息。

### 1.2.3 裂缝长度与深度分析

基于裂缝一般呈现条状特点, 使用裂缝的长度来表示裂缝的大小。在三维空间中, 采用欧式聚类法分割被标记为裂缝的体系, 同时计算其包围盒, 使用包围盒长轴的长度来描述裂缝的长度。

在计算裂缝的深度时, 提取每一个裂缝所在的局部平面, 使用 RANSAC 构造三维平面约束。在裂缝区域的点云中随机选取 3 个点, 计算其初始平面方程  $Ax + By + Cz + D = 0$ 。计算裂缝区域每一个点到初始平面的距离  $d_i = |Ax_i + By_i + Cz_i + D|$ , 若  $d_i$  小于距离阈值(默认采用 0.02m), 则记录该点为模型的内样本点, 根据内样本点计算新的平面方程。迭代每一个点到新的平面方程的距离, 寻找最优内样本点和平面方程。计算裂缝区



域点云到最优平面方程的最大值,记为裂缝的深度信息 $d_m$ 。

2 裂缝识别实验

2.1 Mask R-CNN 数据集与训练

目前关于裂缝的数据集较少,包含裂缝的RGBD数据集则更少。选用 Red Wood<sup>[21]</sup>公开数据集中 HUT 03056 序列作为实验数据集。此序列包含一个完整的木屋结构,墙体中存在部分开裂的情况。同时使用 labelme 标注工具,将序列中的裂缝标注出来以供 Mask R-CNN 网络训练,训练参数如表 1 所示。

表 1 Mask R-CNN 训练参数

Table 1 The training parameters of Mask R-CNN			
迭代/次	网络批量规模	主干网络	损失函数
200	4	ResNet-50	CrossEntropy

2.2 裂缝识别实验结果对比

在 Red Wood 数据集上训练 Mask R-CNN 网络,并使用训练后的网络识别数据集和真实场景中的裂缝,如图 4 所示。

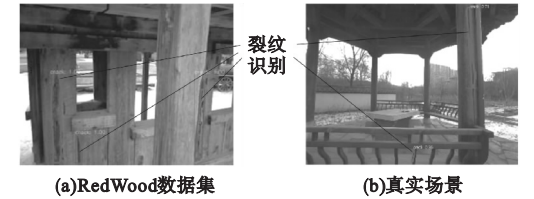


图 4 Mask R-CNN 识别结果

按 1.2.2 方法将融合进三维模型的裂缝信息以体素的形式提取出来,通过式(2)将体素投影到识别的掩码图像中,与所识别的结果拼接。在 Red Wood 数据集和真实场景中均超过了 Mask R-CNN 的识别精度,表明所提出将融合的裂缝信息投影到二维平面来补偿识别的方法是有效的(见表 2)。

表 2 与 Mask R-CNN 识别精度对比

Table 2 The comparison of accuracy between proposed method and Mask R-CNN %		
方法	Red Wood 数据集	真实场景
Mask R-CNN	49.94	32.11
多分辨率重建	88.13	54.30

3 多分辨率三维重建实验

3.1 实验结果对比

图 5 为三维重建部分的实验结果,使用 DirectX 显示当前工作的实时结果,和 Bundle Fusion 一致。

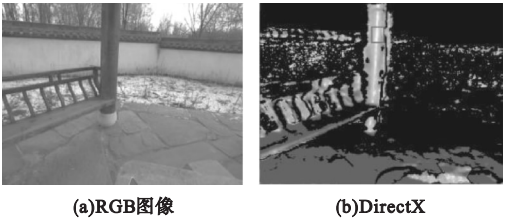


图 5 三维重建运行效果

图 6 为多分辨率三维重建与 Bundle Fusion(BF)重建结果,多分辨率的三维重建将木结构裂缝表面刻画得更清晰。

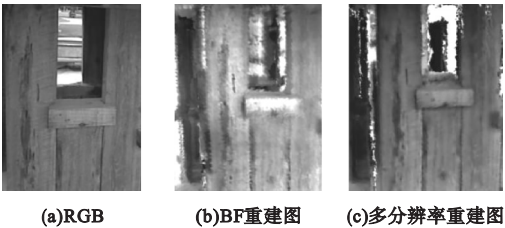


图 6 BF 重建与多分辨率重建结果对比

图 6 BF 重建与多分辨率重建结果对比

3.2 实验结果分析

Bundle Fusion 采用 Voxel Hashing 算法进行建图,其 TSDF 模型是该算法的核心,此模型使用间距 10 mm 的体素表达重建的场景。在这种均匀的体素空间中,精度每提升 1 倍,体素的显存占用量就会扩大 8 倍。将整个空间划分为基本块和细化块。当重建精度为 10 mm 时,Bundle Fusion 需占用 468.75 MB 显存空间,当重建精度是 2.5 mm 时,需占用 30 000 MB 显存的空间。笔者提出的方法仅对部分区域细化后,达到 2.5 mm 的精度仅需占用 3 843.75 MB 显存空间,在保持显存空间不变下可以提高三维重建的精度。

4 裂缝长度与深度分析

4.1 裂缝长度分析实验结果对比

实验中重建和识别裂缝的效果如图 7 所示。在三维体素空间中标注每一个裂缝所在的位置,当体素转为网格后,将裂缝对应体素位置用深色标注,并用方框标志每个裂缝的区域。图 7(a)为 Red Wood 数据集中的重建与裂缝识别结果,场景包含 11 个裂缝,正确识别 10 个裂缝;图 7(b)为真实场景中的重建与裂缝识别结果,场景包含 2 个裂缝,正确识别 2 个裂缝。在没有裂缝的区域,图 7(a)中误识别 1 处,图 7(b)中误识别 2 处。这是由于将图像识别网络与三维重建相结合,会在多次图像识别中造成误判。为解决此问题,采用面积过滤的方法,将所识别的长度小于阈值(0.2 m)的区域舍弃,以此提高识别精度。

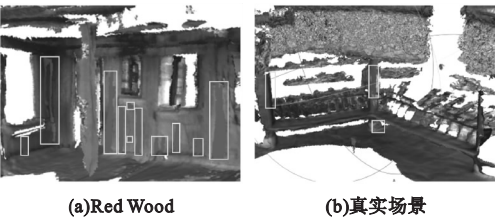


图 7 裂缝识别与重建

**Fig. 7** The recognition and reconstruction of crack

由于 Red Wood 数据集中不包含裂缝的长度信息,因此从 RGB 图中读取裂缝两端像素值,再读取深度图中对应的深度值,即对应像素到相机光心的距离,根据两点坐标计算得出裂缝的长度。表 3 为场景中 9 个裂缝的长度测量值与实际值的对比,其中,裂缝实际长度为现场采用手工测量方式获取。细化前为单一分辨率三维重建计算得到的裂缝长度,其平均误差率为 12.3%,细化后为多分辨率三维重建计算得到的裂缝长度,平均误差率为 9.3%。

4.2 裂缝深度分析实验结果对比

使用 RANSAC 算法拟合每一个裂缝所在平面,并计算裂缝的最大深度。由于 Red

表 3 裂缝长度分析

Table 3 The analysis of crack length

编号	未细化值/m	细化值/m	实际长度/m	R/%	R'/%
Red Wood1	0.862	0.843	0.831	3.7	1.4
Red Wood2	0.621	0.627	0.660	5.9	5.0
Red Wood3	0.633	0.658	0.713	11.2	7.7
Red Wood4	1.178	1.171	1.094	7.7	7.0
Red Wood5	0.288	0.272	0.265	8.7	2.6
Red Wood6	0.472	0.464	0.396	19.2	17.2
Red Wood7	0.904	0.982	1.051	14.0	6.6
Real Scene1	0.751	0.763	0.892	15.8	14.5
Real Scene2	0.314	0.329	0.418	24.9	21.3

注:R 为未细化误差率;R'细化后误差率。

Wood 数据集仅包含每一帧像素到光心的距离,很难通过像素到光心的距离得出裂缝的实际深度值,因此裂缝深度分析结果仅与真实场景中人工测量值进行比对(见表 4)。

表 4 裂缝深度分析

Table 4 The analysis of cracks depth

编号	未细化值/m	细化值/m	实际深度/m	R/%	R'/%
Real Scene1	0.048	0.040	0.037	29.7	8.1
Real Scene2	0.009	0.011	0.016	43.8	31.2

经计算,单一分辨率三维重建的平均深度误差率为 36.8%,多分辨率三维重建的平均深度误差率为 19.7%。

5 结 论

- (1)通过细化裂缝处重建效果,在 Red Wood 数据集中将图像识别精度从原 49.94% 提高到 88.13%,在真实场景中将图像识别精度从原 32.11% 提高到 54.30%。
- (2)与人工测量结果比对计算误差率,在 Reg Wood 数据集中,与单一分辨率重建方法相比,多分辨率方法的裂缝长度平均误差率从 12.3% 降低到 9.3%。
- (3)在真实场景中,与单一分辨率重建方法相比,多分辨率方法的深度平均误差率从 36.8% 降低到 19.7%。

## 参考文献

- [1] CHEN Z Y, ZHU E C, PAN J L, et al. Structural performance of typical beam-column joints in Yingxian wood pagoda-an experimental study [J]. Key engineering materials, 2012, 517:669-676.
- [2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [J]. Proceedings of the IEEE conference on computer vision and pattern recognition, 2014:580-587.
- [3] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. Advances in neural information processing systems, 2015, 28.
- [4] GIRSHICK R. Fast R-CNN [C]. Proceedings of the IEEE international conference on computer vision. [S. l.]: [s. n.], 2015: 1440-1448.
- [5] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]. Proceedings of the IEEE international conference on computer vision. [S. l.]: [s. n.], 2017:2961-2969.
- [6] 马健, 阎卫东, 刘国奇. 基于 YOLO v5 的古建筑木结构裂缝检测方法[J]. 沈阳建筑大学学报(自然科学版), 2021, 37(5):927-934.  
(MA Jian, YAN Weidong, LIU Guoqi. Crack detection method of ancient building wood structure based on YOLO v5 [J]. Journal of Shenyang jianzhu university (natural science), 2021, 37(5):927-934.)
- [7] 韩士伟, 李一挥. 面向古建筑的倾斜摄影与精细化建模[J]. 测绘地理信息, 2020, 45(4):46-50.  
(HAN Shiwei, LI Yihui. Oblique photography and fine modeling for ancient buildings [J]. Journal of geomatics, 2020, 45(4):46-50.)
- [8] 李艳, 宿勇军. 三维激光扫描技术在古建筑园林建模与系统开发中的应用[J]. 测绘地理信息, 2018, 43(2):35-37.  
(LI Yan, SU Yongjun. Application of 3D laser scanning technology in modeling and system development of ancient architectural gardens [J]. Journal of geomatics, 2018, 43(2):35-37.)
- [9] 韦虎, 张丽艳, 张辉. 双目立体测量中多视角深度图同时融合算法[J]. 计算机辅助设计与图形学学报, 2008, 20(11):7.  
(WEI Hu, ZHANG Liyan, ZHANG Hui. Simultaneous fusion algorithm of multi-view depth map in binocular stereo measurement [J]. Journal of computer-aided design and computer graphics, 2008, 20(11):7.)
- [10] 陈姝丽, 王涛, 凌王章. 基于多视角影像的古建筑三维模型重建[J]. 科技风, 2020, No. 410(6):207-208.  
(CHEN Shuli, WANG Tao, LING Wangzhang. Reconstruction of ancient architecture 3D model based on multi-view image [J]. Technology wind, 2020, 410(6):207-208.)
- [11] 阮競芸, 厉犇. 多视角深度图配准算法下破损古建筑三维重建[J]. 计算机仿真, 2021, 38(7):157-160.  
(RUAN Jingyun, LI Ben. 3D reconstruction of damaged ancient buildings based on multi-view depth map registration algorithm [J]. Computer simulation, 2021, 38(7):157-160.)
- [12] 李务军, 吴斌. 基于 Kinect 深度图像的三维重建[J]. 微型机与应用, 2016, 35(5):55-57.  
(LI Wujun, WU Bin. 3D reconstruction of depth image based on Kinect [J]. Microcomputer and applications, 2016, 35(5):55-57.)
- [13] 陈占军, 林姚宇, 龚咏喜, 等. 古建筑视觉三维重建系统设计与实现[J]. 计算机应用与软件, 2021, 38(8):17-22.  
(CHEN Zhanjun, LIN Yaoyu, GONG Yongxi, et al. Design and implementation of 3D visual reconstruction system for ancient buildings [J]. Computer applications and software, 2021, 38(8):17-22.)
- [14] IZADI S, KIM D, HILLIGES O, et al. Kinect fusion: real-time 3D reconstruction and interaction using a moving depth camera [C]// Proceedings of the 24th annual ACM symposium on user interface software and technology. [S. l.]: [s. n.], 2011:559-568.
- [15] DAI A, NIEßNER M, ZOLLHÖFER M, et al. Bundle fusion: real-time globally consistent 3D reconstruction using on-the-fly surface reintegration [J]. ACM transactions on graphics, 2017, 36(4):1-19.
- [16] NIEßNER M, ZOLLHÖFER M, IZADI S, et al. Real-time 3D reconstruction at scale using voxel hashing [J]. ACM transactions on graphics, 2013, 32(6):1-11.
- [17] MCCORMAC J, HANDA A, DAVISON A, et al. Semantic fusion: dense 3D semantic mapping with convolutional neural networks [C]. International conference on robotics and automation. IEEE [s. n.], 2017:4628-4635.
- [18] XU B, LI W, TZOUMANIKAS D, et al. Mid-fusion: octree-based object-level multi-instance dynamic slam [C]. International conference on robotics and automation. IEEE, 2019:5231-5237.
- [19] WU W, GUO L, GAO H, et al. YOLO-SLAM: a semantic SLAM system towards dynamic environment with geometric constraint [J]. Neural computing and applications, 2022, (1):6011-6026.
- [20] NG P C, HENIKOFF S. SIFT: predicting amino acid changes that affect protein function [J]. Nucleic acids research, 2003, 31(13):3812-3814.
- [21] CHOI S, ZHOU Q Y, MILLER S, et al. A large dataset of object scans, 10.48550/arXiv.1602.02481 [P]. 2016.  
(责任编辑:王国业 英文审校:范丽婷)