

第38卷 第X期
2025年X月

中 国 公 路 学 报
China J. Highw. Transp.

Vol. 38 No. X
X. 2025

引文格式：侯佳林,侯榕榕,鲍跃全.基于三维高斯溅射的桥梁自动化三维重建方法[J].中国公路学报,2025,38(X):000-000.

HOU Jia-lin, HOU Rong-rong, BAO Yue-quan. Automated 3D Reconstruction of Bridges Based on 3D Gaussian Splatting [J]. China Journal of Highway and Transport, 2025, 38 (X): 000-000.

文章编号:1001-7372(2025)00-0001-12

基于三维高斯溅射的桥梁自动化三维重建方法

侯佳林^{1,2},侯榕榕^{*1,2},鲍跃全^{1,2}

(1. 哈尔滨工业大学 土木工程智能防灾减灾工信部重点实验室,黑龙江 哈尔滨 150090;
2. 哈尔滨工业大学 土木工程学院,黑龙江 哈尔滨 150090)

摘要:数字孪生与元宇宙对于实现桥梁智慧运维起着至关重要的作用,其中桥梁自动化三维重建是构建数字孪生模型及元宇宙平台的关键。针对现有桥梁三维重建方法难以实现几何-语义一体化重建的问题,提出一种基于三维高斯溅射(3D Gaussian Splatting, 3DGS)的桥梁高保真自动化三维重建方法。首先,针对桥梁构件区域边界特征,提出基于分割一切模型(Segment Anything Model, SAM)的桥梁语义分割模型,对无人机拍摄的桥梁多视角图像进行高精度语义分割,生成桥面、索塔等构件语义掩码。随后,在场景表征的三维高斯核函数中增加语义特征属性,实现复杂几何和语义特征的自适应表征。最后,将三维高斯投影至二维图像空间,利用可微光栅化算法生成语义特征图和渲染图,通过几何-表观-语义多属性联合优化,实现桥梁几何结构保真度与语义关联精确性相统一的三维重建。本文以广东省南沙大桥为例进行验证,试验结果表明,重建的桥梁三维模型可以高保真还原桥梁纹理细节及语义信息,三维重建的平均交并比为87.2%,总体准确率为91.6%。本文所提出方法有效解决了传统方法中几何与语义信息割裂问题,为桥梁全生命周期数字孪生提供兼具物理精确性与语义可解释性的三维基础模型。

关键词:桥梁工程;三维重建;三维高斯溅射;悬索桥;语义分割

中图分类号:U445.6 文献标志码:A

Automated 3D Reconstruction of Bridges Based on 3D Gaussian Splatting

HOU Jia-lin^{1,2}, HOU Rong-rong^{*1,2}, BAO Yue-quan^{1,2}

(1. Key Laboratory of Intelligent Disaster Prevention and Mitigation in Civil Engineering of the Ministry of Industry and Information Technology, Harbin Institute of Technology, Harbin 150090, Heilongjiang, China; 2. School of Civil Engineering, Harbin Institute of Technology, Harbin 150090, Heilongjiang, China)

Abstract: Digital twin and metaverse play a crucial role in realizing the intelligent operation and maintenance of bridges, in which the automated 3D reconstruction of bridges is the key to build the digital twin model and metaverse platform. To address the challenge of integrating geometric and semantic reconstruction in existing 3D bridge reconstruction methods, this study proposed a high-fidelity automated 3D bridge reconstruction method based on 3D Gaussian splatting (3DGS). First, to capture the boundary features of bridge components, a semantic segmentation

收稿日期:2024-12-31

基金项目:国家自然科学基金项目(52208310)

作者简介:侯佳林(2001-),男,辽宁锦州人,工学博士研究生,E-mail:23S133114@stu.hit.edu.cn。

* 通讯作者:侯榕榕(1991-),女,辽宁铁岭人,教授,博士研究生导师,工学博士,E-mail:rrhou@hit.edu.cn。

model for bridges was proposed using the Segment Anything Model (SAM), enabling high-precision semantic segmentation of multi-view UAV-captured bridge images and generating semantic masks for components such as decks and bridge towers. Next, semantic feature attributes were incorporated into the 3D Gaussian kernel functions of scene representation, achieving adaptive characterization of complex geometric and semantic features. Finally, 3D Gaussians were projected into 2D image space, and differentiable rasterization algorithms were utilized to generate semantic feature maps and rendered images. Through joint optimization of geometric, appearance, and semantic attributes, a unified 3D reconstruction was realized, balancing geometric fidelity and semantic accuracy. In this paper, Nansha Bridge in Guangdong Province is used as an example for validation, and the experimental results show that the reconstructed 3D bridge model achieves high-fidelity restoration of texture details and semantic information, with a mean Intersection over Union (mIoU) of 87.2% and an overall accuracy of 91.6%. The proposed method effectively resolves geometric detail loss and semantic fragmentation in traditional approaches, providing a 3D foundational model with both physical precision and semantic interpretability for lifecycle digital twins of bridges.

Keywords: bridge engineering; 3D reconstruction; 3D Gaussian splatting; suspension bridge; semantic segmentation

Received 31 Dec. 2024

Funding: National Natural Science Foundation of China (52208310)

* Corresponding author. E-mail address: rrhou@hit.edu.cn.

0 引言

大跨度桥梁作为重要的交通基础设施,其长期安全运行对于社会经济的稳定发展至关重要。近年来,数字孪生与元宇宙技术的迅速发展,为桥梁智慧运维提供了新思路^[1]。数字孪生通过物理实体与虚拟模型的实时交互,为桥梁全生命周期运营维护提供了高效智能的解决方案^[2]。三维重建作为构建数字孪生模型的核心,在桥梁设计、建造、运营、养护等阶段均发挥着关键作用^[3-5]。融合语义信息的高精度三维重建,能够准确获取桥梁各构件的几何及语义信息,从而服务于桥梁长期健康监测和运营维护^[6-7]。然而,现有三维重建方法^[8-10]通常仅关注表观及几何建模,无法准确提取桥梁各构件的语义信息,难以满足桥梁数字孪生和智慧运维的实际需求。因此,如何在桥梁三维模型中有效融入语义信息,成为实现桥梁数字孪生及运维智能化的关键。

近年来,基于激光雷达扫描和无人机倾斜摄影的三维重建方法取得了显著进展^[11-14]。相较于基于激光雷达扫描点云的方法,基于图像的三维重建方法具有数据采集简单、成本低廉等显著优势。通过消费级无人机获取高分辨率图像,结合结构恢复运

动(Structure from Motion, SfM)^[15]和多视图立体视觉(Multi-view Stereo, MVS)^[16]方法,生成高精度三维点云模型。Xu 等^[17]提出一种基于分布式多视图的桥梁三维重建方法,首先基于平面先验辅助的 MVS 算法计算桥梁子模型,再通过分布式计算将其整合为完整桥梁模型,方法重建精度可达厘米级。Morgenthaler 等^[18]利用无人机图像结合 SfM 生成桥梁三维模型,用于桥梁状态评估。Pan 等^[19]提出一种半自动化桥梁结构表面模型生成算法,利用分类树算法结合几何特征实现构件识别,随后通过泊松表面重建生成模型,召回率和精度均超过 0.8。然而,上述方法在细粒度语义表达和自动化建模方面仍存在局限。因此,如何实现桥梁三维语义模型的高效高精度重建是目前亟需解决的问题。

目前,桥梁构件识别方法主要可分为自下而上、自上而下和基于深度学习三类^[20]。自下而上的方法根据法线、曲率、点密度等局部几何特征,对提取的点云数据进行聚类和分割,从而实现桥梁构件的准确识别。Truong-Hong 等^[21]提出一种基于单元和体素的三段式区域生长方法用于桥梁构件识别,其点云评估 F1 得分高达 0.93。Mehranfar 等^[22]结合传统描述符和模糊 C 均值算法,实现桥墩、桥台、

主体和栏杆的准确语义分割。Xia 等^[23]开发了一种为典型桥梁定制的多尺度局部描述符,能够准确描述其几何形状和拓扑关系。然而,上述方法计算复杂,耗时较长,且对噪声和离群点敏感,因此限制了其在实际工程中的应用。与自下而上的方法不同,自上而下的方法基于桥梁构件的几何特征和设计规则等领域知识,实现桥梁构件的分割与识别。Lu 等^[20]首次提出基于桥梁整体框架分解的构件自动识别方法,首先通过递归切片算法构建桥面和桥墩组件,再基于几何特征和密度直方图进行局部构件分割。Zhao 等^[24]利用模板匹配方法识别桥梁构件的位置和类型。Yan 等^[25]采用增量式横截面对齐方法提取钢梁桥构件语义信息。Qin 等^[26]将点云数据沿竖向划分为若干区间,基于点云密度变化和投影滤波算法实现桥墩上部结构、下部结构和地面的准确分割。然而,这一类方法在处理形状复杂或几何特征模糊的构件时表现不佳,且过于依赖先验假设,难以灵活适应实际桥梁所涉及的多样化场景。

近年来,随着人工智能的快速发展,深度学习在桥梁工程领域取得了成功应用^[27-29]。PointNet^[30]及其改进版本 PointNet++^[31]作为典型的端到端点云处理模型,开创性地将深度学习方法应用于点云分类任务,实现了点云全局特征提取与局部几何建模的高效融合。Kim 等^[32]提出一种基于 PointNet 的桥梁构件自动识别方法,实现了桥面、桥墩与背景的高效分类。Yang 等^[33-34]开发一种基于加权超点图模型的深度学习方法,用于桥梁构件精确分类。Hu 等^[35]提出一种结合递归二叉树网络的斜拉桥结构三维语义模型重建方法,成功实现桥梁构件

及其拓扑关系的准确建模。然而,上述方法均依赖大量训练数据,标注成本高,同时在桥梁纹理信息表达和复杂构件特征捕捉方面存在明显局限。

针对上述问题,本文提出一种桥梁自动化三维语义模型重建方法。首先,通过对图像分割模型^[36]进行微调,实现了桥梁构件的高精度语义分割。随后,基于三维高斯溅射(3D Gaussian splatting, 3DGS)^[37],利用可微光栅化算法将语义信息映射至三维高斯模型,实现二维语义信息与三维高斯模型的深度融合。本研究有效解决现有方法存在的人工标注点云繁琐、语义分割效率低下以及语义信息表达能力不足等挑战,建立了基于多视角图像的桥梁三维语义模型重建通用框架。试验结果表明,本文所提出方法能够实现桥梁三维模型的几何-语义一体化高效重建,为桥梁数字孪生提供数字模型基础。

1 研究方法

本研究所提出的桥梁三维重建方法流程如图 1 所示。首先,利用微调后的图像分割模型对无人机采集的多视角图像进行语义分割,得到高精度语义掩码。同时,通过 SfM 获取相机位姿和初始点云,为三维高斯提供初始几何和空间位置信息。随后,为场景中每个三维高斯赋予一个初始化语义特征属性,并将三维高斯投影至二维图像空间,利用可微光栅化算法生成相应的语义特征图和渲染图。其中,语义特征图基于语义一致性损失进行优化,而渲染图则通过对比与真实图像的重建损失实现优化。最终,获得一个包含丰富语义信息桥梁高保真三维模型。

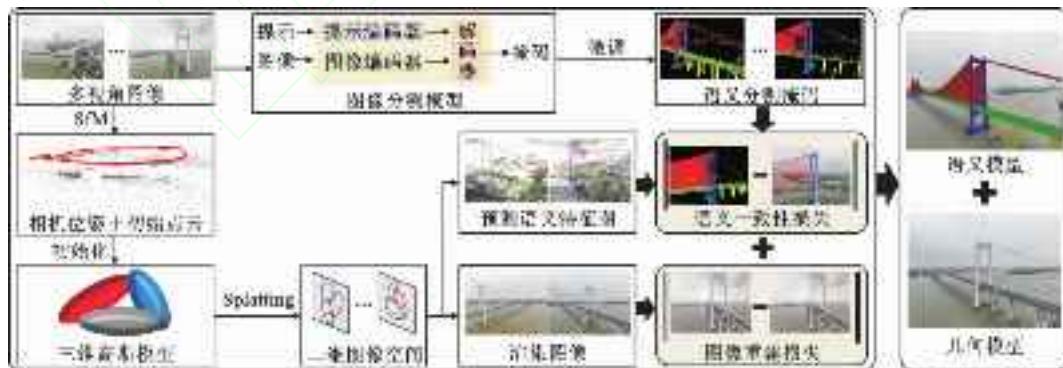


图 1 桥梁三维重建流程

Fig. 1 Flowchart of Bridge 3D Reconstruction

1.1 桥梁语义分割

1.1.1 图像分割模型

分割一切模型(Segment Anything Model, SAM)^[36]是一种通用图像分割框架,能够分割图像

上的任何目标。如图 2 所示,SAM 模型主要由图像编码器(Image Encode)、提示编码器(Prompt Encode)和掩码解码器(Mask Encode)三部分组成,支持多种形式的输入,如点、边界框和文本。然而,



图 2 SAM 网络架构

Fig. 2 SAM Architecture

SAM 在桥梁工程领域的特定分割任务中存在局限性。具体而言,桥梁语义分割任务需要对每个像素进行精确分类以生成多类别预测掩码,而 SAM 仅能在提示输入的引导下生成若干二值化掩码,缺乏明确的类别信息。如图 3 所示,基于 SAM 的桥梁图像分割结果中同一类别的像素未能正确分类,从而难以获取正确语义信息。此外,SAM 所生成的掩码数量随提示变化,难以满足桥梁语义分割对固定多类别掩码输出的需求。因此,为了提升 SAM 在桥梁语义分割任务中的适用性,需要对其进行有针对性的微调,使其能够精准感知区域边界并准确识别语义类别。



图 3 基于 SAM 的桥梁图像语义分割掩码

Fig. 3 Semantic Segmentation Masks for Bridge Images Based on SAM

1.1.2 基于微调 SAM 的语义分割模型

SAM 的图像编码器在 SA-1B 数据集上进行预训练,涵盖丰富的自然场景语义信息,因此其参数迭代良好,能够有效适应下游任务。为了利用 SAM 图像编码器的强大特征提取能力,本文保留基于视觉转换器(Vision Transformer, ViT)构建的图像编码器原始架构,在训练过程中冻结其原始参数,仅对插入的轻量级适配器进行微调。所提出模型的整体架构如图 4 所示。

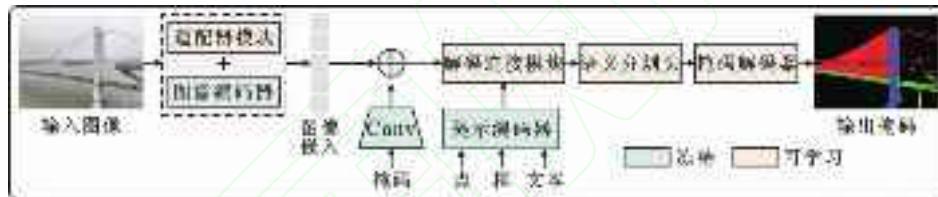


图 4 微调 SAM 整体架构图

Fig. 4 Overall Architecture of Fine-Tuned SAM

在微调过程中,适配器被插入到 ViT 的每个 Transformer 模块中,以增强模型的特征适应能力。每个 Transformer 模块由多头自注意力层和多层次感知机层组成。假设输入的原始特征为 $X \in \mathbb{R}^{M \times n}$,其经过自注意力层后输出为 $X' \in \mathbb{R}^{M \times n}$

$$X' = \text{Attention}(Q, K, V) = \text{Softmax}(QK^T / \sqrt{d_k})V \quad (1)$$

式中: Q, K, V 分别为特征 $X \in \mathbb{R}^{M \times n}$ 的查询、键和值向量; d_k 为自注意力维度。

为了调整特征空间以适应桥梁语义分割任务,在自注意力层之后引入轻量级适配器模块。如图 5 所示,适配器由一个降维全连接层(Down)、ReLU 激活函数和升维全连接层(Up)构成

$$\text{Adapter}(X) = W_{\text{up}} \text{ReLU}(W_{\text{down}} X) \quad (2)$$

式中: $W_{\text{up}} \in \mathbb{R}^{n \times l}$ 为升维矩阵; $W_{\text{down}} \in \mathbb{R}^{l \times n}$ 为降维矩阵。

在图像嵌入特征经过编码器处理后,掩码连接模块将图像嵌入信息与提示编码器的输出特征进行融合,并结合上下文语义信息和位置提示信息,对融

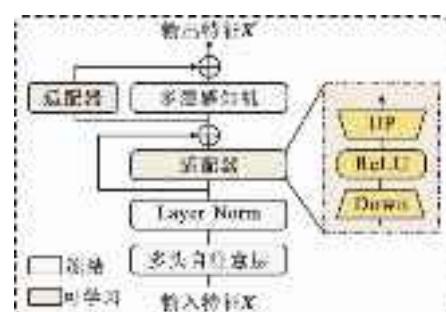


图 5 适配器插入 Transformer 模块架构

合后的特征进行进一步编码处理。随后,设计一个语义分割头进行输入特征解码,将解码后的特征接入分类器。分类器将每个像素的特征映射至类别空间,输出每个像素的类别预测概率。最后,采用如下交叉熵损失函数计算分类结果与真实标签之间的误差

$$L_{\text{cross}} = - \sum_{q=1}^O y_q \log(\hat{y}_q) + (1 - y_q) \log(1 - \hat{y}_q) \quad (3)$$

式中: \hat{y}_q 为分类器输出; y_q 为真实标签; O 为类别数。交叉熵损失通过最小化预测结果与真实标签之间的差异,能够在训练过程中有效处理类别不平衡问题,从而实现每个类别像素的正确分类。最终,所提出基于 SAM 的微调模型能够输出准确的语义分割结果,为后续三维重建提供可靠的语义输入。

1.2 基于 3DGS 的桥梁三维模型重建

1.2.1 3DGS 基本原理

由于实时高质量渲染及场景可编辑性等优势,3DGS^[37]在三维重建领域引起广泛关注。3DGS 显式建模场景的辐射场,在保留体渲染细节的同时,通过 CUDA 加速的可微光栅化技术实现高效实时渲染。首先,将无人机采集的多视角图像作为输入,利用 SfM 提取稀疏点云,这些点云提供了初始场景的几何信息。随后,根据提取的稀疏点云初始化生成三维高斯,其中每个三维高斯 $G(x)$ 由世界坐标系中以均值点 μ 为中心的三维协方差矩阵 Σ 表示

$$G(x) = \exp\left[-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right] \quad (4)$$

式中: x 为三维高斯在场景中的任意位置。三维协方差矩阵 Σ 由缩放矩阵 S 和旋转矩阵 R 构建,以确保其为半正定

$$\Sigma = RSS^T R^T \quad (5)$$

与传统依赖计算资源密集型的光线步进技术进行体积渲染的方法不同^[38],3DGS 采用基于瓦片的光栅化技术实现高效场景渲染,将图像划分为 16×16 的图像块,并剔除位于视椎体外的三维高斯。然后,对与图像块相交的二维高斯按照投影深度进行排序,并通过透明度混合计算每个图像块内像素的最终颜色

$$C(x') = \sum_{i=1}^N c_i \alpha'_i \prod_{j=1}^{i-1} (1 - \alpha'_j) \quad (6)$$

式中: x' 为查询的像素位置; N 为与该像素关联的二维高斯按深度排序后的集合; c_i 表示由球谐系数生成的第 i 个三维高斯投影到二维图像的颜色; α'_i 为第 i 个三维高斯在二维图像中的最终不透明度,由三维高斯的不透明度与其在二维投影中的密度相乘得到; $j(1 \leq j \leq i-1)$ 为累积透射率计算中的迭代索引。

1.2.2 三维高斯语义特征初始化

原始三维高斯仅通过几何属性以及外观特征表征场景,缺乏对场景进行精细化表征的语义信息。为了提升 3DGS 对于三维场景的表征能力,本研究为每个三维高斯赋予一个语义特征向量 $s \in \mathbb{R}^K$,其

中维度 K 为语义类别数量,每个维度 $s^{(r)}$ 表示该三维高斯所在区域属于语义类别 r 的概率。在初始化阶段,每个三维高斯的语义特征向量 $s \in \mathbb{R}^K$ 随机生成,假设其服从标准正态分布。这种随机初始化方法确保所有三维高斯的语义特征在初始时为均匀分布,允许模型在训练过程中通过学习逐步调整和优化语义特征。因此,本研究中每个三维高斯 G 的属性通过以下五元组表示

$$G = (\mu, \Sigma, \alpha, c, s) \quad (7)$$

1.2.3 基于可微光栅化的语义映射

为了实现高效的语义信息映射,本文利用可微光栅化算法将三维高斯投影至二维图像空间,生成相应的语义特征图。三维高斯的语义信息能够在图像空间中进行有效传递,并支持梯度反向传播,从而实现三维高斯语义特征的优化。通过相机位姿和投影矩阵,将每个三维高斯从世界坐标系转换到相机坐标系,给定视图变换矩阵 W 和三维协方差矩阵 Σ ,二维协方差矩阵 Σ' 由下式计算

$$\Sigma' = JW\Sigma W^T J^T \quad (8)$$

式中: J 为投影变换仿射近似的雅可比矩阵。

在图像平面上,与三维高斯的其他属性类似,每个像素点的语义特征值 $F(u)$ 通过所有投影高斯的语义特征向量加权求和计算得到

$$F(u) = \sum_{i=1}^N T(u) \alpha_i s_i \quad (9)$$

式中: $T(u)$ 为三维高斯在像素 u 处的透射率; s_i 为第 i 个三维高斯的语义特征向量。

由于光栅化过程是可微的,在后续优化过程中,每个三维高斯的语义特征能够通过梯度反向传播更新,从而提高其在三维空间中语义信息的准确性。上述利用可微光栅化计算语义特征图并通过反向传播优化语义特征的过程,使得最终生成的三维高斯能够准确捕捉场景中的语义特征,从而提升模型的语义表达能力。

1.2.4 三维高斯特征优化

原始 3DGS 通过计算渲染图像和真实图像之间的光度和结构相似性误差,进行三维高斯几何 Σ 、颜色 c 和透明度 α 等属性的优化,其损失函数定义如下

$$L = (1 - \lambda)L_{\text{image}} + \lambda L_{\text{D-SSIM}} \quad (10)$$

式中: λ 为权重系数; L_{image} 用于量化像素级的颜色误差; $L_{\text{D-SSIM}}$ 侧重捕捉图像的结构相似性,从而提升模型在感知质量上的表现。

为了提高三维高斯语义特征的准确性,本研究

引入语义一致性损失。基于交叉熵计算生成的语义特征图与语义分割模型生成的语义掩码之间的差异,语义一致性损失定义为

$$L_{\text{semantic}} = - \sum_{u \in I} \sum_{r=1}^K Z_r(u) \log(\hat{Z}_r(u)) \quad (11)$$

式中: $Z_r(u)$ 为像素 u 的真实语义标签; $\hat{Z}_r(u)$ 为生成的语义特征图中类别 r 的预测值。

为了同时优化三维高斯的几何和语义特征,使得模型在渲染质量和语义精度之间取得良好平衡,本研究结合语义一致性损失和原始3DGS的渲染损失定义如下损失函数

$$L_{\text{total}} = (1 - \lambda_1)L_{\text{image}} + \lambda_1 L_{\text{D-SSIM}} + \lambda_2 L_{\text{semantic}} \quad (12)$$

式中: λ_1 和 λ_2 分别为平衡图像重建损失和语义一致性损失的权重系数。

2 试验数据集制备

由于目前尚无开源的桥梁语义分割图像数据集,本文使用搭载高清镜头的消费级无人机对桥梁进行全方位、多角度拍摄,以获取高分辨率多视角图像。采集完成后,对图像数据进行语义标注与预处理,建立可供语义分割模型训练的数据集。本试验共采集570张原始图像,并从中随机选取500张进行语义标注,为后续模型训练和验证提供可靠数据基础。

2.1 数据采集与预处理

南沙大桥位于中国广东省珠江三角洲中部,是一座横跨珠江的公路悬索桥,也是珠江三角洲高速公路网的重要组成部分。大桥主跨长1 200 m,两侧跨各为400 m,由两座193 m高的索塔和8个基础支撑结构组成。试验过程中,采用配备高清相机的消费级无人机从多个视角捕捉桥梁图像。无人机通过内置的定点巡航功能,预先规划飞行路线,以确保采集数据能够全面覆盖桥梁结构,并提升采集效率。试验所采用无人机与相机参数如表1所示。

表1 无人机和相机参数

Table 1 Parameters of the UAV and Camera

设备	项目	参数
DJI Mavic 3 Classic	尺寸	347.5 mm×283 mm×107.7 mm
	起飞重量	895 g
	最大可倾斜角度	35°
	GPS 悬停精度	水平±0.3 m, 垂直±0.1 m
4/3 CMOS Hasselblad	有效像素	2 000万
	光圈范围	f/2.8~f/11
	等效焦距	24 mm
	最大照片尺寸	5 280 像素×3 956 像素

在图像采集过程中,无人机与桥梁结构保持约10 m距离,以确保拍摄角度的清晰性和覆盖范围。所采集的图像分辨率为5 280像素×3 956像素,部分图像样本如图6(a)所示。随后,通过SfM技术和COLMAP提供的MVS管道^[39],从采集图像中生成稀疏点云并提取相机参数。生成的稀疏点云与相机位姿如图6(b)所示,其中黑色点代表点云数据,反映了桥梁的几何形态,红色块则表示相机的外部参数,包括空间位置和姿态信息。将上述获取的稀疏点云和相机参数作为三维先验,为后续三维重建提供立体视觉支持。此外,为保证重建模型的绝对尺度准确性,无人机在图像采集时同步记录GPS坐标与高度信息(水平精度±0.3 m,垂直精度±0.1 m),并在COLMAP稀疏重建阶段作为先验约束优化相机参数与点云尺度,确保重建模型与真实场景的绝对比例一致。

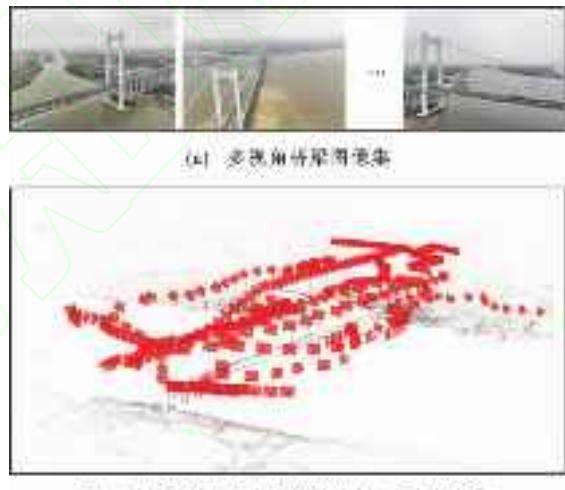


图6 无人机采集数据及数据预处理结果

Fig. 6 Data Acquisition by UAV and Data Preprocessing Results

2.2 语义标签标注

为了实现由输入图像到桥梁语义分割掩码的有监督学习,需要对桥梁各构件区域进行逐像素标注。本文使用图像标注工具Labelme^[40]为每个像素点分配标签,以明确该像素点属于桥面、主缆、索塔或基础。图7展示了无人机拍摄的桥梁原始图像及其所对应的语义标签图。

3 试验验证与结果分析

为验证本文所提出的方法在实际工程中的有效性,本文以南沙大桥为对象,开展桥梁三维几何-语义一体化重建的试验验证。首先,将所采集到的图

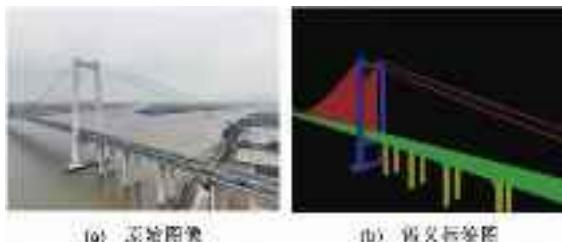


图 7 原始图像与其对应的语义标签图

Fig. 7 Original Image and Its Corresponding Semantic Label Map

像数据及其对应的真实语义掩码按照一定比例划分为训练集、验证集和测试集, 用于微调语义分割模型。随后, 将原始图像数据和生成的稀疏点云作为输入, 并结合语义分割模型生成的语义掩码, 进一步训练三维语义高斯模型, 最终实现自动化三维语义模型重建。

3.1 试验环境

本文模型训练过程的硬件环境为英特尔 Core i9-14900K 处理器和 NVIDIA GeForce RTX 4090 显卡, 内存为 64 GB。试验操作系统为 Windows 11, 采用开源的 PyTorch 2.5.0 深度学习框架构建网络模型, CUDA 版本为 12.6。编程语言使用 Python, 版本为 3.12.4。

3.2 模型训练与验证

3.2.1 语义分割模型训练

如第 1.1 节所述, 本文通过标注的桥梁语义分割图像对 SAM 模型进行微调。在训练过程中, 将标注的 500 张语义图像随机划分为训练集、验证集和测试集, 其中 80% 的图像用于训练, 其余各 10% 的图像分别用于验证和测试。所有输入图像的尺寸统一调整为 1 024 像素 \times 1 024 像素, 并通过仿射变换、随机翻转和随机旋转操作进行数据增强, 以提升数据的多样性和数量, 进而提高模型的泛化能力。

训练过程中使用随机梯度下降优化器进行梯度更新, 并采用余弦退火学习率调度策略以动态调整学习率。其中, 初始学习率设置为 1×10^{-5} , 动量参数设置为 0.9, 权重衰减参数为 1×10^{-5} 。训练批次大小设置为 32, 训练总轮次设定为 800 轮。图 8 显示了训练过程中训练集、验证集和测试集损失值变化趋势。从图中可以看出, 模型在训练初期损失值快速下降, 随后逐渐趋于平稳, 表明模型在逐步适应数据特征。在经过约 400 轮训练后, 损失值下降速度明显放缓, 并最终在约 600 轮后趋于稳定。训练集和验证集的损失值均收敛于 0.12 附近, 表明模型

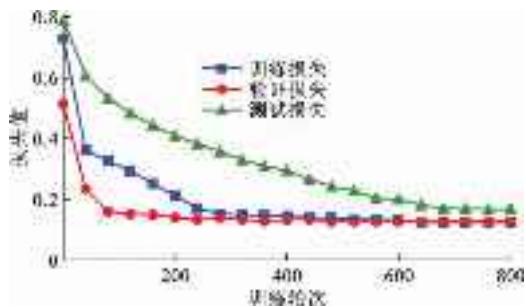


图 8 损失函数曲线

Fig. 8 Loss Function Curve

在经过充分训练后已达到较优性能, 为后续三维语义高斯模型训练提供精准语义信息支持。测试集的损失曲线收敛趋势与训练集和验证集一致, 最终损失值稳定在 0.16 左右。尽管测试集包含未参与训练的样本, 但其损失值与验证集接近, 表明模型具有较好的泛化能力。

3.2.2 三维语义高斯模型训练

三维语义高斯模型训练的目的是为了将语义分割模型得到的语义掩码通过可学习的语义特征映射至三维空间, 从而实现桥梁三维几何-语义模型重建。首先, 将采集到的多视角图像下采样至原来的四分之一并校正图像畸变。随后, 使用 Adam 优化器训练三维语义高斯模型, 学习率设置为 0.01, 并进行 30 000 次迭代以确保模型充分收敛。语义特征向量初始化为 128 维, 在训练过程中压缩至 32 维以减少内存开销。在损失权重设置上, 为在图像重建和语义一致性之间取得良好平衡, 将二者的权重系数分别设置为 0.8 和 0.2。对于分辨率为 1 320 像素 \times 989 像素的场景, 模型训练时间约为 35 min, 内存占用约 16 GB。图 9 展示了三维语义高斯的可视化结果。

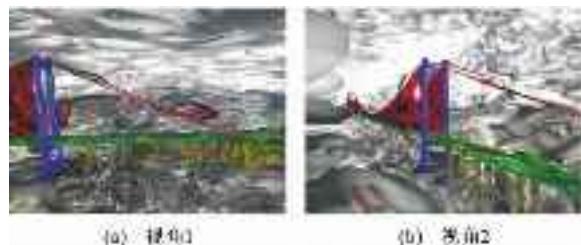


图 9 不同视角下三维语义高斯可视化

Fig. 9 Visualization of Semantic 3D Gaussian from Different Viewpoints

3.3 模型评估与分析

为了定量评估本文所提出的方法对于桥梁三维重建的有效性, 本文从图像重建质量、模型几何精度和语义分割精度 3 个方面进行综合评估。图像重建

质量采用峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、结构相似指数(Structural Similarity, SSIM)和学习感知图像斑块相似度(Learned Perceptual Image Patch Similarity, LPIPS)衡量。其中,PSNR用于衡量渲染图像与原始图像之间的差异,PSNR值越大表示图像重建质量越高,SSIM用于评估渲染图像与原始图像在结构上的相似性,值越大表明模型对图像几何结构的保真度越高,LPIPS则用于衡量两张图像在视觉感知上的相似性,值越小表示图像在感知上的表现越优越。模型几何精度通过准确性(Accuracy, Acc)、完整性(Completeness, Comp)、综合性能(Overall)及F1 Score等指标进行量化评估。准确性定义为重建点云与真值点云之间的平均距离,值越小表明几何形态还原越精准,完整性表示真值点云中被正确重建的比例,值越大表明模型覆盖范围越完整,综合性能为Acc与Comp的调和平均数,F1 Score则为两者的综合平衡指标。此外,本文从渲染语义图像和点云语义分割两个方面来评估模型的语义分割性能,采用总体准确率(Overall Accuracy, OA)、交并比(Intersection-over-Union, IoU)和平均交并比(Mean Intersection-over-Union, mIoU)作为衡量指标,计算公式如下

$$A_{OA} = \sum_{i=0}^{N-1} P_{ii} / \left(\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_{ij} \right) \quad (13)$$

$$I_{IoU} = P_{ii} / \left(\sum_{j=0}^{N-1} P_{ij} + \sum_{i=0}^{N-1} P_{ji} - P_{ii} \right) \quad (14)$$

$$I_{mIoU} = \frac{1}{N} \sum_{i=0}^{N-1} P_{ii} / \left(\sum_{i=0}^{N-1} P_{ij} + \sum_{j=0}^{N-1} P_{ji} - P_{ii} \right) \quad (15)$$

式中: P_{ii} 为类别*i*上正确分类的像素点数; P_{ij} 为实际类别为*i*但被预测为*j*的像素点数。 A_{OA} 反映被正确分割的点云在总点云数中的比例, I_{IoU} 用于衡量单个语义类别的分割结果与真实标签的重叠程度, I_{mIoU} 则表示所有语义类别的分割结果与真实标签重叠程度的平均值。

3.3.1 图像重建质量评估

本文所提出方法在三维语义高斯模型的训练过程中引入语义特征,使三维高斯在优化过程中不仅关注桥梁各构件的几何和表观属性,还能够捕捉各构件间的语义关联。这种语义约束不仅有助于减少模型在结构细节处的模糊,还能够进一步优化三维高斯的分布,使模型能够更准确地还原构件的外观特征和层次关系,从而提升重建质量。如图10所示,本文模型能够清晰重建索塔表面裂纹、主缆防护



图 10 桥梁损伤可视化

Fig. 10 Visualization of Bridge Damage

层破损及基础混凝土剥落等表观损伤,验证了本文方法在纹理细节还原上的高保真性。

相较于传统人工巡检,本文方法利用多视角图像自动化重建桥梁三维模型,通过模型中几何-语义信息一体化表达,巡检人员可快速在模型中确定损伤区域的空间位置和范围,相比传统现场逐一巡检的方式,显著提升了巡检效率。此外,重建模型中各构件的语义信息可直接用于自动标记构件病害位置,实现损伤区域快速定位,从而极大地提高了后续桥梁健康监测与维护的效率。

如表2所示,本文所提出方法在PSNR和SSIM指标上均优于原始3DGS,表明其在渲染质量和保真度上表现更为优越。此外,相较于原始3DGS本文所提出方法的LPIPS值较低,表明其能够更好地保持图像的感知一致性,有效解决重建过程中的失真问题。

表 2 渲染图像质量对比

Table 2 Comparison of Rendering Image Quality

方法	PSNR ↑	SSIM ↑	LPIPS ↓
原始3DGS	29.18	0.822	0.132
本文方法	32.46	0.833	0.126

3.3.2 模型几何精度评估

由于目前尚无开源的桥梁点云数据集,本文从3D Warehouse数据集中选取典型斜拉桥网格模型,其几何复杂度与南沙大桥主跨结构相似,可有效模拟实际工程场景。如图11(a)所示,通过Blender软件对该网格模型进行随机离散化处理作为真值点云。基于Blender内置渲染管线,模拟无人机多视角采集条件,生成包含300张图像的多视角数据集。为确保仿真试验中点云模型的尺度精度,本文在Blender环境中通过内置路径规划与相机控制模块生成精确的相机参数(包括焦距、位置和姿态),从而为稀疏重建提供绝对尺度约束。

基于MVS和本文方法的三维重建结果如图11

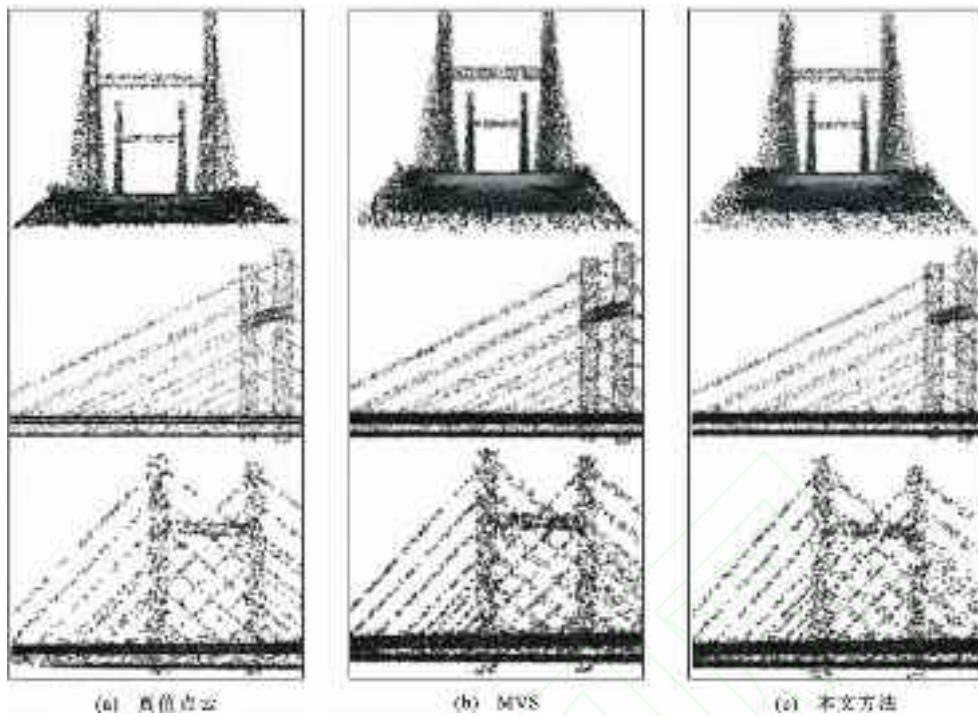


图 11 重建点云模型

Fig. 11 Reconstructed Point Cloud Model

(b)、(c) 所示。相较于基于真实场景实拍图像的 MVS 重建结果, 图 11(b) 中 MVS 方法生成的点云较为稀疏。这主要是由于本文所使用的多视角图像来源于 Blender 仿真渲染, 桥梁纹理较为均匀且缺乏背景, 导致在 MVS 的像素级深度估计过程中可用于特征匹配的纹理信息不足, 从而生成的点云密度相对较低。不同于传统 MVS 方法基于像素级深度估计获得稠密点云的方式, 本文方法在 COLMAP 生成的稀疏点云基础上, 通过基于图像重建误差的梯度反馈, 动态优化三维高斯的空间位置和形状参数。同时, 利用语义一致性损失进一步约束三维高斯的分布, 使其更准确地贴合桥梁各构件的边界, 从而实现几何与语义的高精度联合建模。

为了定量评估重建模型的几何精度, 本文采用迭代最近点算法 (ICP) 配准重建点云与真值点云, 计算准确性 (Acc)、完整性 (Comp)、综合性能 (Overall) 及 F1 Score 指标。如表 3 所示, 本文方法重建模型的准确性为 1.27 cm 优于 MVS 方法, 表明其能够更精准地还原桥梁几何形态。然而, MVS 方法在主缆等狭长构件区域具有更高的点云密度与完整性。在综合平衡指标方面, 本文方法重建模型的 F1 Score 为 92.31% 与 MVS 方法接近, 表明其在保证几何精度的同时, 能够较完整重建桥面、索塔等主体结构。需要指出的是, 该仿真实验主要旨在评估

表 3 点云重建质量对比

Table 3 Comparison of Point Cloud Reconstruction Quality

方法	Acc/cm	Comp/cm	Overall	F1 Score
MVS	1.36	0.92	1.14	94.78
本文方法	1.27	1.83	1.55	92.31

本文方法在理想化场景下的重建精度上限。在实际应用场景中, 受图像质量、光照条件变化以及环境遮挡等复杂因素影响, 重建模型几何精度可能产生一定程度的衰减。

传统基于 MVS 的稠密重建方法依赖密集点云实现高精度建模, 但存在计算资源消耗过大的问题。本文方法通过稀疏高斯分布表征场景, 结合基于 CUDA 加速的光栅化技术, 在保证几何精度的同时实现模型的高效重建, 内存占用仅为 16 GB, 从而支持动态环境下模型的快速更新, 适合集成至轻量化运维平台。此外, MVS 方法仅能生成几何模型, 需依赖后处理提取构件语义信息, 流程繁琐且效率低下。然而, 本文方法通过联合优化几何与语义特征, 能够实现几何-语义信息一体化高精度建模, 为桥梁数字孪生与智慧运维提供更高效的解决方案。

3.3.3 渲染语义图像评估

将重建的三维语义模型渲染至对应视角的二维图像, 计算其与真实掩码之间的 IoU 和 mIoU 以评估模型的语义分割性能。如表 4 所示, 模型对桥面、

表 4 渲染语义图像分割结果

Table 4 Segmentation Results of Rendering Semantic Image

构件类别	IoU/%	mIoU/%
桥面	89.6	88.5
主缆	81.5	
索塔	90.3	
基础	92.5	

主缆、索塔和基础的 IoU 分别为 89.6%、81.5%、90.3% 和 92.5%。其中,由于主缆在悬索桥结构中尺度较小,且与背景易产生混淆,导致主缆的识别准确率偏低。所有结构构件的平均交并比为 88.5%,表明模型在悬索桥语义分割任务中具有良好的精度,能够有效识别桥梁构件的不同类别。为了直观评价本文所提出方法的语义分割效果,图 12 展示了悬索桥在不同视角下渲染得到的语义分割图。从图中可以看出,模型能够准确分割桥面、主缆、索塔及

基础等构件,分割边界清晰,语义分割精度较高。

3.3.4 点云语义分割评估

为了进一步量化本文模型在三维空间中的语义分割精度,验证语义信息从二维图像到三维空间映射的准确性,将本文方法输出的语义点云与人工标注的真值语义点云进行对比。真值点云是基于原始三维高斯模型输出的稀疏点云,这种稀疏表征方式可显著降低内存占用与计算复杂度。通过开源点云处理软件 CloudCompare 逐点标注桥面、主缆、索塔及基础,标注过程中参考无人机多视角图像中桥面与吊杆的锚固边界、主缆与索塔的垂吊连接边界、桥面与基础的过渡边界等构件分界特征,结合 CloudCompare 的剖面切割工具实现精细化分割。由于原始三维高斯点云与本文方法输出的语义点云数量级一致,不存在密度差异对计算精度的影响,保证了对比的公平性。

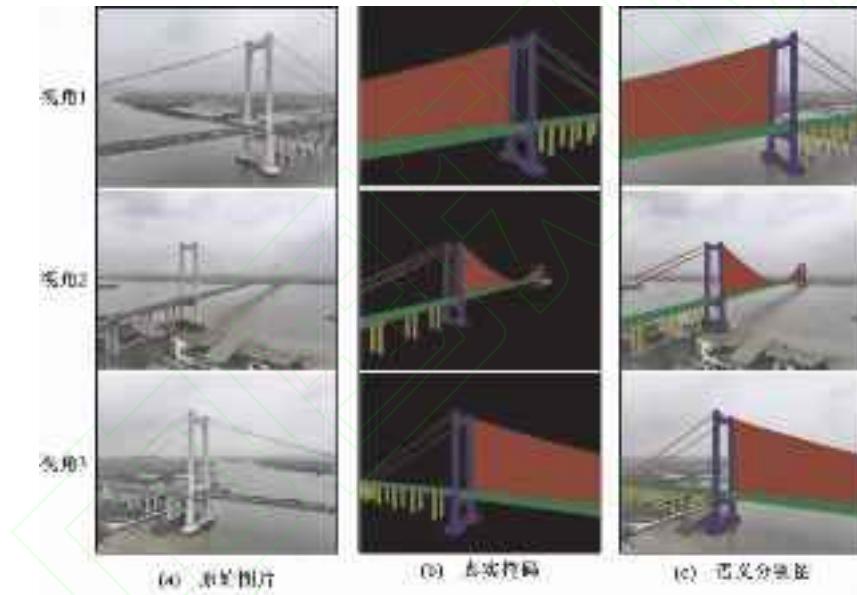


图 12 不同视角下三维语义模型渲染结果对比

Fig. 12 Comparison of 3D Semantic Model Rendering Results from Different Perspectives

试验结果三维点云的 mIoU 为 87.2%,表明本文所提出方法能够在复杂三维场景中准确区分桥梁主体的不同构件。总体准确率 OA 达到 91.6%,验证本文所提出方法在三维空间中的语义表达能力,及其良好的鲁棒性与稳定性。点云语义分割可视化结果如图 13 所示,从对比图可以看出,本文方法重建模型在桥面、索塔和基础部分的分割表现良好,重建的语义点云与真实语义点云具有较高的一致性。然而,由于语义分割掩码中主缆区域与背景较难区分,导致主缆部分点云分割效果不够理想,出现一定程度的漏分和误分现象。

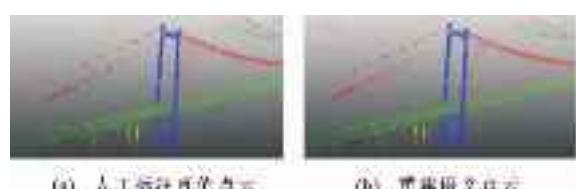


图 13 点云语义分割可视化

Fig. 13 Visualization of Point Cloud Semantic Segmentation

4 结语

本文提出了一种基于三维高斯溅射的桥梁三维语义模型重建方法,通过微调图像分割模型对无人

机采集的多视角图像进行高精度语义分割,实现了二维语义信息与三维高斯模型的深度融合,成功重建了具有丰富语义信息的高精度桥梁三维模型。主要结论如下:

(1)本文通过微调 SAM 图像分割模型,提出了一种适用于桥梁结构的语义分割模型,实现了对桥梁图像的高精度语义分割,为三维高斯模型训练提供了可靠的语义输入。

(2)通过引入三维高斯语义特征属性,基于可微光栅化算法,本文实现了从二维图像语义信息到三维语义高斯模型的精准映射,从而避免了对三维点云的人工标注。

(3)在三维高斯优化过程中,本文所提出的方法通过捕捉各构件间的语义关联,为三维高斯的几何及表观属性优化提供有效约束,不仅有效解决了传统三维建模中语义表达能力不足的问题,同时提升了几何模型精度及渲染质量。

(4)以广东南沙大桥为研究对象,试验结果表明本文提出的方法在图像重建质量和语义分割精度方面均取得了良好的表现。其中,渲染图像 PSNR 和 SSIM 指标分别为 32.46 和 0.833,LPIPS 值低至 0.126;渲染得到的语义图像在桥面、主缆、索塔和基础部分的 IoU 分别达到 89.6%、81.5%、90.3% 和 92.5%,三维语义点云 mIoU 为 87.2%,总体准确率为 91.6%,充分验证了本文方法的有效性。

作为桥梁数字孪生系统的数字模型基础,本文重建得到的模型能够高保真地还原桥梁主体构件的几何形状及纹理细节,从而真实反映桥面、主缆等构件的表观损伤,大幅减少人工巡检的工作量。此外,模型中的语义信息能够服务于无人机巡检航线规划,并为后续各构件损伤的定量分析提供依据。然而,本文方法在吊索类细长构件的几何重建精度方面仍存在不足,未来将探索针对提升此类构件重建效果的优化策略。

参考文献:

References:

- [1] FRANCIOSI M, KASSER M, VIVIANI M. Digital twins in bridge engineering for streamlined maintenance and enhanced sustainability [J]. Automation in Construction, 2024, 168: 105834.
- [2] YANG Y L, ZHU Y C, CAI C S. Research progress and prospect of digital twin in bridge engineering [J]. Advances in Structural Engineering, 2024, 27 (2): 333-352.
- [3] JIANG F, DING Y, SONG Y, et al. An architecture of life-cycle fatigue management of steel bridges driven by Digital Twin [J]. Structural Monitoring and Maintenance, 2021, 8 (2): 187-201.
- [4] YASIN YİĞİT A, UYSAL M. Virtual reality visualisation of automatic crack detection for bridge inspection from 3D digital twin generated by UAV photogrammetry [J]. Measurement, 2025, 242: 115931.
- [5] GUO X Y, FANG S G. Digital twin based lifecycle modeling and state evaluation of cable-stayed bridges [J]. Engineering with Computers, 2024, 40 (2): 885-899.
- [6] 王凌波,王秋玲,朱 钊,等.桥梁健康监测技术研究现状及展望[J].中国公路学报,2021,34(12):25-45.
WANG Ling-bo, WANG Qiu-ling, ZHU Zhao, et al. Current status and prospects of research on bridge health monitoring technology [J]. China Journal of Highway and Transport, 2021, 34 (12): 25-45.
- [7] 孙利民,尚志强,夏 煜.大数据背景下的桥梁结构健康监测研究现状与展望[J].中国公路学报,2019,32(11):1-20.
SUN Li-min, SHANG Zhi-qiang, XIA Ye. Development and prospect of bridge structural health monitoring in the context of big data [J]. China Journal of Highway and Transport, 2019, 32 (11): 1-20.
- [8] YAN Y J, HAJJAR J F. Geometric models from laser scanning data for superstructure components of steel girder bridges [J]. Automation in Construction, 2022, 142: 104484.
- [9] WANG F, ZOU Y, DEL REY CASTILLO E, et al. Automated UAV path-planning for high-quality photogrammetric 3D bridge reconstruction [J]. Structure and Infrastructure Engineering, 2024, 20 (10): 1595-1614.
- [10] 马 威,涂 强,潘建平,等.桥梁实景三维高斯辐射场建模[J].测绘学报,2024,53(9):1694-1705.
MA Wei, TU Qiang, PAN Jian-ping, et al. 3D Gaussian radiation field modeling for real-scene bridges [J]. Acta Geodaetica et Cartographica Sinica, 2024, 53 (9): 1694-1705.
- [11] HAN Y T, FENG D M, WU W W, et al. Geometric shape measurement and its application in bridge construction based on UAV and terrestrial laser scanner [J]. Automation in Construction, 2023, 151: 104880.
- [12] 余加勇,薛现凯,陈昌富,等.基于无人机倾斜摄影的公路边坡三维重建与灾害识别方法[J].中国公路学报,2022,35(4):77-86.
YU Jia-yong, XUE Xian-kai, CHEN Chang-fu, et al. Three-dimensional reconstruction and disaster identification of highway slope using unmanned aerial vehicle-based oblique photography technique [J]. China Journal of Highway and Transport, 2022, 35 (4): 77-86.
- [13] LIU Y F, NIE X, FAN J S, et al. Image-based crack assessment of bridge piers using unmanned aerial vehicles and three-dimensional scene reconstruction [J]. Computer-Aided Civil and Infrastructure Engineering, 2020, 35 (5): 511-529.
- [14] 王羽尘,于 斌,陈晓阳,等.基于激光雷达点云的道路几何信息提取与数字化建模研究[J].中国公路学报,2023,36(3):45-60.
WANG Yu-chen, YU Bin, CHEN Xiao-yang, et al. Extraction and digital modeling of road geometric information using

- LiDAR data point clouds [J]. China Journal of Highway and Transport, 2023, 36 (3): 45-60.
- [15] SNAVELY N, SEITZ S M, SZELISKI R. Photo tourism: Exploring photo collections in 3D [M] // Seminal Graphics Papers: Pushing the Boundaries, Volume 2. New York, NY, USA: ACM, 2023: 515-526.
- [16] HIEP V H, KERIVEN R, LABATUT P, et al. Towards high-resolution large-scale multi-view stereo [C] // 2009 IEEE Conference on Computer Vision and Pattern Recognition. June 20-25, 2009, Miami, FL, USA. IEEE, 2009: 1430-1437.
- [17] XU Y, ZHANG J. UAV-based bridge geometric shape measurement using automatic bridge component detection and distributed multi-view reconstruction [J]. Automation in Construction, 2022, 140: 104376.
- [18] MORGENTHAL G, HALLERMANN N, KERSTEN J, et al. Framework for automated UAS-based structural condition assessment of bridges [J]. Automation in Construction, 2019, 97: 77-95.
- [19] PAN Y, DONG Y Q, WANG D L, et al. Three-dimensional reconstruction of structural surface model of heritage bridges using UAV-based photogrammetric point clouds [J]. Remote Sensing, 2019, 11 (10): 1204.
- [20] LU R D, BRILAKIS I, MIDDLETON C R. Detection of structural components in point clouds of existing RC bridges [J]. Computer-Aided Civil and Infrastructure Engineering, 2019, 34 (3): 191-212.
- [21] TRUONG-HONG L, LINDBERGH R. Automatically extracting surfaces of reinforced concrete bridges from terrestrial laser scanning point clouds [J]. Automation in Construction, 2022, 135: 104127.
- [22] MEHRANFAR M, AREFI H, ALIDOOST F. Knowledge-based 3D reconstruction of bridge structures using UAV-based photogrammetric point cloud [J]. Journal of Applied Remote Sensing, 2021, 15 (4): 044503.
- [23] XIA T, YANG J, CHEN L. Automated semantic segmentation of bridge point cloud based on local descriptor and machine learning [J]. Automation in Construction, 2022, 133: 103992.
- [24] ZHAO Y-P, WU H T, VELA P A. Top-down partitioning of reinforced concrete bridge components [C] // Computing in Civil Engineering 2019. Atlanta, Georgia. American Society of Civil Engineers, 2019: 275-283.
- [25] YAN Y J, HAJJAR J F. Automated extraction of structural elements in steel girder bridges from laser point clouds [J]. Automation in Construction, 2021, 125: 103582.
- [26] QIN G C, ZHOU Y, HU K X, et al. Automated reconstruction of parametric BIM for bridge based on terrestrial laser scanning data [J]. Advances in Civil Engineering, 2021, 2021 (1): 8899323.
- [27] 鲍跃全,李惠.人工智能时代的土木工程[J].土木工程学报,2019,52(5):1-11.
BAO Yue-quan, LI Hui. Artificial intelligence for civil engineering [J]. China Civil Engineering Journal, 2019, 52 (5): 1-11.
- [28] SUN L M, SHANG Z Q, XIA Y, et al. Review of bridge structural health monitoring aided by big data and artificial intelligence: From condition assessment to damage detection [J]. Journal of Structural Engineering, 2020, 146 (5): 04020073.
- [29] 邓露,褚鸿鹄,龙砾芝,等.基于深度学习的土木基础设施裂缝检测综述[J].中国公路学报,2023,36(2):1-21.
DENG Lu, CHU Hong-hu, LONG Li-zhi, et al. Review of deep learning-based crack detection for civil infrastructures [J]. China Journal of Highway and Transport, 2023, 36 (2): 1-21.
- [30] CHARLES R Q, HAO S, MO K C, et al. PointNet: Deep learning on point sets for 3D classification and segmentation [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 77-85.
- [31] QI C R, YI L, SU H, et al. PointNet++: Deep hierarchical feature learning on point sets in a metric space [EB/OL]. 2017: 1706. 02413. <https://arxiv.org/abs/1706.02413v1>.
- [32] KIM H, YOON J, SIM S H. Automated bridge component recognition from point clouds using deep learning [J]. Structural Control and Health Monitoring, 2020, 27 (9): e2591.
- [33] YANG X F, DEL REY CASTILLO E, ZOU Y, et al. Automated semantic segmentation of bridge components from large-scale point clouds using a weighted superpoint graph [J]. Automation in Construction, 2022, 142: 104519.
- [34] YANG X F, DEL REY CASTILLO E, ZOU Y, et al. Semantic segmentation of bridge point clouds with a synthetic data augmentation strategy and graph-structured deep metric learning [J]. Automation in Construction, 2023, 150: 104838.
- [35] HU F Q, ZHAO J, HUANG Y, et al. Structure-aware 3D reconstruction for cable-stayed bridges: A learning-based method [J]. Computer-Aided Civil and Infrastructure Engineering, 2021, 36 (1): 89-108.
- [36] KIRILLOV A, MINTUN E, RAVI N, et al. Segment anything [C] // 2023 IEEE/CVF International Conference on Computer Vision (ICCV). October 1-6, 2023, Paris, France. IEEE, 2023: 3992-4003.
- [37] KERBL B, KOPANAS G, LEIMKUEHLER T, et al. 3D Gaussian splatting for real-time radiance field rendering [J]. ACM Transactions on Graphics, 2023, 42 (4): 1-14.
- [38] MILDENHALL B, SRINIVASAN P P, TANCIK M, et al. NeRF: Representing scenes as neural radiance fields for view synthesis [C] // Computer Vision-ECCV 2020. Cham: Springer, 2020: 405-421.
- [39] SCHÖNBERGER J L, FRAHM J M. Structure-from-motion revisited [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 4104-4113.
- [40] RUSSELL B C, TORRALBA A, MURPHY K P, et al. LabelMe: A database and web-based tool for image annotation [J]. International Journal of Computer Vision, 2008, 77 (1): 157-173.

附录 A

基于 3DGS 的桥梁三维语义重建算法伪代码

算法: 基于 3DGS 的桥梁三维重建

输入: 多视角图像集 $I = \{I_k\}_{k=1}^K$

输出: 三维高斯语义模型 $G = \{G_n\}_{n=1}^M$

步骤一: 数据准备

将多视角图像集 I 输入 COLMAP, 获取初始稀疏点云 $D = \{d_n\}_{n=1}^M$ 和相机参数;

基于微调后的语义分割模型, 对每张图片 I_k 生成真实语义标签集合 $\{Z_k\}_{k=1}^K$

步骤二: 初始化三维高斯

初始化三维高斯集合 $G = \emptyset$

for $d_n \in D$ **do:**

 将点的位置初始化为三维高斯的中心 $\mu_n \leftarrow d_n$

 协方差矩阵初始化为各向同性 $\Sigma_n \leftarrow R S S^T R^T$

 基于图像插值初始化颜色 c_n

 透明度初始化为固定值 $\alpha_n \leftarrow \alpha_{init}$

 随机初始化语义特征向量 $s_n \leftarrow s_n \sim N(0, 1)$

 将 $G_n = (\mu_n, \Sigma_n, \alpha_n, c_n, s_n)$ 加入 G

end for

步骤三: 可微光栅化计算

for $I_k \in I$ **do:**

 获取视图变换矩阵 $W_k \leftarrow$ 相机位姿

for $G_n \in G$ **do:**

 计算投影后的协方差矩阵 $\Sigma'_n \leftarrow J W_k \Sigma_n W_k^T J^T$

 渲染投影后的三维高斯 $G'_n \leftarrow (\mu_n, \Sigma')$

end for

 累加渲染结果 $\tilde{I}_k \leftarrow \sum_n G'_n$

 提取语义特征图 $\hat{Z}_k \leftarrow \tilde{I}_k$

end for

步骤四: 迭代优化

for $t = 1$ **to** 最大迭代次数 **do:**

for $I_k \in I$ **do:**

 计算图像重建损失 $\leftarrow L_{image}(I_k, \tilde{I}_k)$

 计算结构相似性损失 $\leftarrow L_{DSSIM}(I_k, \tilde{I}_k)$

 计算语义一致性损失 $\leftarrow L_{semantic}(Z_k, \hat{Z}_k)$

end for

 梯度计算 $\nabla L \leftarrow \partial L / \partial$

 参数更新 $(\mu_n, \Sigma_n, \alpha_n, c_n, s_n) \leftarrow Adam(\nabla L)$

end for

return G