Journal of Image and Signal Processing 图像与信号处理, 2026, 15(4), 593-604
Published Online October 2026 in Hans. https://www.hanspub.org/journal/jisp
https://doi.org/10.12677/jisp.2026.154052

Hans汉斯

基于三维高斯泼溅的港口场景高效轻量化三维重建

陈张扬[1]，王雪雯[1*]，胡剑秋[2]，诸　明[1]

[1]江苏海洋大学理学院，江苏 连云港
[2]中国船舶集团有限公司第七一六研究所，江苏 连云港

收稿日期：2026年9月27日；录用日期：2026年10月19日；发布日期：2026年10月29日

摘　要

针对智慧港口数字孪生建设对高精度、轻量化三维重建的需求，现有激光雷达、倾斜摄影和神经辐射场等传统方法在港口场景中存在的成本高、建模慢、动态适配差等瓶颈，提出一种基于三维高斯泼溅的港口场景三维重建方法。构建涵盖稀疏视角鲁棒建模与轻量化部署的三维高斯泼溅重建框架，采用改进运动恢复结构初始化、港口结构约束的损失函数和分级自适应密度控制等关键技术。在港口场景的实验表明，该方法单场景建模耗时7.1 min，较基线仅增加4.4%，较基础3DGS在重建精度、渲染效率及轻量化部署方面优势显著，可有效支撑智慧港口数字孪生建设对高精度、轻量化三维重建的需求。

关键词

三维重建，3D高斯泼溅，智慧港口，渐进式分辨率训练，低显存训练

High-Efficiency Lightweight 3D Reconstruction of Port Scenes Based on 3D Gaussian Splatting

Zhangyang Chen[1], Xuewen Wang[1*], Jianqiu Hu[2], Ming Zhu[1]

[1]School of Science, Jiangsu Ocean University, Lianyungang Jiangsu
[2]No. 716 Research Institute, CSSC, Lianyungang Jiangsu

Received: September 27, 2026; accepted: October 19, 2026; published: October 29, 2026

Abstract

Aiming at the requirements of high-precision, high-efficiency and lightweight 3D reconstruction for

*通讯作者。

文章引用: 陈张扬, 王雪雯, 胡剑秋, 诸明. 基于三维高斯泼溅的港口场景高效轻量化三维重建[J]. 图像与信号处理, 2026, 15(4): 593-604. DOI: 10.12677/jisp.2026.154052

smart port digital twin construction, this paper analyzes the bottlenecks of traditional methods such as LiDAR, oblique photography and NeRF in port scenes, including high cost, slow modeling and poor dynamic adaptability. A port scene 3D reconstruction method based on 3D Gaussian Splatting is proposed. A 3D Gaussian Splatting reconstruction framework covering sparse-view robust modeling and lightweight deployment is constructed. Key technologies including improved Structure-from-Motion initialization, multi-constraint loss function with port structural constraints, and hierarchical adaptive density control are adopted. Experiments on port scenes show that the single-scene modeling time is only 7.1 min, only 4.4% more than the baseline. Compared with baseline 3DGS, the proposed method has significant advantages in reconstruction accuracy, rendering efficiency and lightweight deployment, which can effectively support high-precision and lightweight 3D reconstruction for smart port digital twin construction.

Keywords

3D Reconstruction, 3D Gaussian Splatting, Smart Port, Progressive Resolution Training, Low-Video-Memory Training

Copyright © 2026 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

当前全球港口正加速数字化转型，数字孪生、智能调度、安全监测等业务均需高精度、高效率、实时渲染的三维模型支撑。现有港口三维重建技术存在明显短板：激光雷达成本高昂、建模周期漫长[1]；倾斜摄影难以还原岸桥、集装箱等异形钢结构细节[2]；神经辐射场(NeRF) [3]训练耗时久、渲染卡顿、动态物体建模效果差[4]，无法适配港口全天候动态作业环境。2023 年 Kerbl 等提出 3D 高斯泼溅(3DGS) [5]算法，依托显式高斯表达与可微光栅化渲染优势，凭借建模快、细节好、实时交互强等特点成为三维重建前沿热点，国内外学者已围绕基础算法优化展开研究[6] [7]，并初步应用于测绘、矿山[8]、地质标本[9]等行业。多源数据融合方法[10]虽可提升重建完整性，但数据配准难度大、工程落地难。然而，面向港口多设备、多动态目标、多源异构数据混杂的复杂工况，针对性算法改良与落地研究仍较匮乏，缺少成熟的定制化重建体系。

本文以典型港口建筑场景为研究对象，重点模型训练效率低与显存占用等关键难题。港口场景具有占地广阔、要素混杂(集装箱群、装卸机械、船舶、护栏、水域交错分布)、钢制结构强光反射、纹理高度重复、通航与作业带来全天候动态变化、沿海多雾阴雨逆光拍摄[11]等特征，重建难度高。结合智慧港口数字孪生建设[12]需求，确定硬性指标约束：追求厘米级几何精度、单场景建模耗时数分钟以内、渲染帧率高、严控显存与内存占用、保障重建模型的结构完整性与视觉真实感。

本文立足港口场景独有特征，构建基于三维高斯泼溅的港口场景高效重建与低显存训练框架。主要创新点如下：(1) 渐进式分辨率与误差引导密度控制(Progressive-Guided Error-Guided Control，简称 PG-EGC)，通过“先粗后细”渐进式分辨率训练、渲染误差热力图引导的智能密度控制与自适应球谐阶数提升三者协同，在相同训练时间内实现 PSNR 提升。(2) 动态显存池化与梯度检查点(Dynamic Memory Gradient Checkpoint，简称 DM-GC)，通过梯度检查点重计算换显存、动态高斯池化按重要性淘汰、优化器状态压缩以 float16 存储减半显存、分块渲染降低单次渲染峰值四管齐下，在训练时间增加不足 5%前提下将显存占用降低。

2. 港口三维重建关键技术

2.1. 三维高斯泼溅框架

三维高斯泼溅技术重建过程包含五个核心环节：① SfM 点云初始化与 3D 高斯椭球构建、② 投影变换、③ 快速可微光栅化、④ 损失计算与梯度反向传播、⑤ 自适应密度控制。各环节原理与本文具体工作对应关系见图 1。

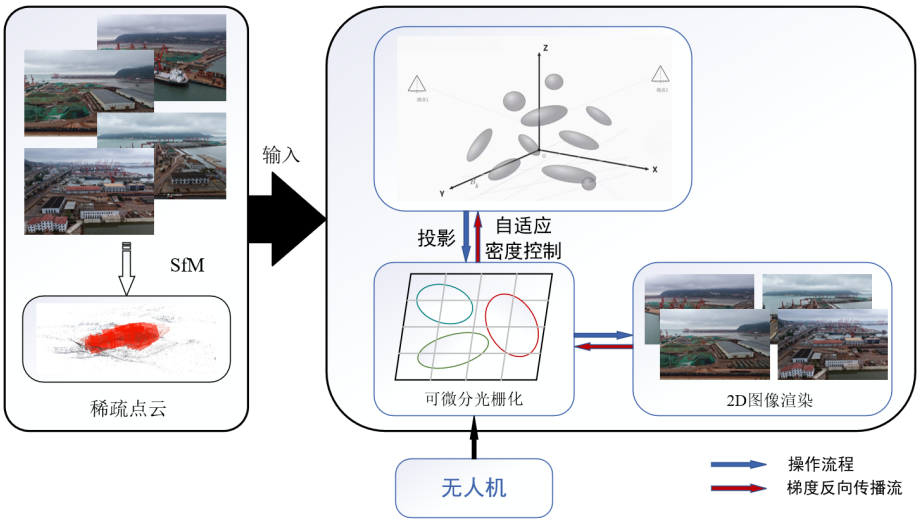


Figure 1. Gaussian splatting flowchart
图 1. 高斯泼溅流程图

环节① SfM 点云初始化与 3D 高斯椭球构建。利用 COLMAP 运动恢复结构算法对多视角图像解算相机位姿与稀疏点云，以点云坐标初始化高斯均值 $\mu \in R^3$。三维高斯分布是 3DGS 场景建模的基础数学原理，其概率密度函数如式(1)所示，能在三维空间生成椭球体以具象化表达实景物体的空间几何轮廓。

三维高斯分布的概率密度函数定义为：

$$f(x)=\frac{1}{(2\pi)^{3/2}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)\right), \tag{1}$$

式中：x 为三维空间坐标；μ 确定高斯基元空间中心；Σ 为协方差矩阵，约束椭球尺寸、外形与朝向，拆解为 $\Sigma = RSS^{\mathrm{T}}$ (R 为旋转矩阵，S 为缩放矩阵)，旋转参数采用四元数[13]表示以保证正交约束。

环节② 投影变换。高斯基元经仿射变换从三维空间映射至二维成像平面，仿射变换保持高斯分布数学特性。因相机成像属非线性透视投影，通过雅可比矩阵 J 提取投影函数偏导数，将局部非线性映射近似为线性变换，简化三维到二维的投影运算量。该环节对应本文 2.2 节 COLMAP 位姿解算与点云初始化流程。

环节③ 快速可微光栅化。渲染采用图像分块并行策略，将画面划分为像素区块，按深度排序后进行 Alpha 透明度分层混合，借助 GPU 多线程并行加速。球谐函数编码不同观测角度下的颜色光照信息，低阶系数刻画漫反射，高阶参数记录镜面高光细节，精准还原集装箱钢板反光、水面波光等港口复杂视觉特征。像素颜色值 C 通过深度顺序和不透明度 α 合成[14]。

环节④ 损失计算与梯度反向传播。预测图像与实拍原图计算 L_1 损失与 SSIM 结构相似性损失，损失梯度沿渲染链路反向传导至每个高斯基元的位置、尺度、球谐颜色、不透明度等参数，结合随机梯度下降自动迭代修正参数，实现端到端优化。

环节⑤ 自适应密度控制。根据梯度幅值动态调整高斯密度：对梯度较大的区域执行克隆或分裂增加基元密度以细化结构，对不透明度过低的区域执行剪枝移除冗余基元。该反馈机制贯穿训练全程，确保算力精准投放至细节丰富区域。本文在此基础上进一步提出 PG-EGC 渐进式误差引导密度控制(1.2 节)与 DM-GC 动态显存池化(1.3 节)，分别从质量优化与显存压缩两个维度改进环节⑤的密度控制策略。

2.2. 渐进式分辨率与误差引导密度控制

针对标准 3DGS 在固定多轮训练中存在“算力分配不均”的问题，本文提出渐进式分辨率与误差引导密度控制方法 PG-EGC。该方法将训练过程划分为三个阶段：骨架期 0~10,000 轮在 1/2 分辨率下以 0 阶球谐函数快速收敛全局几何；过渡期 10,000~15,000 轮切换至全分辨率并解冻 1 阶球谐函数引入漫反射；精细期 15,000~30,000 轮启用 3 阶球谐函数还原高光反射，并结合渲染误差热力图实施差异化密度控制。

渐进式分辨率训练方面，骨架期在 1/2 分辨率下训练，单轮计算量仅为全分辨率的 1/4，加速比公式为 $t_{1/2}=\frac{t_{full}}{4}$ (其中 t_{full} 为全分辨率单轮训练时间)，前 10,000 轮等效节省 7500 轮全分辨率计算预算，节省的算力全部投入到精细期的精细化训练中。低分辨率下高频细节被自然模糊，梯度信号主要由全局几何与本色主导，高斯位置和协方差快速收敛至正确位置。过渡期切换至全分辨率时，对球谐系数做零填充实现平滑过渡，避免剧烈梯度波动。

误差引导密度控制方面，每 100 轮在全分辨率下渲染训练视角，计算逐像素 L_1 误差 $E(x,y)=\left|I_{pred}(x,y)-I_{gt}(x,y)\right|$，经 5×5 均值滤波后归一化至[0, 1]区间，记为 $\hat{E}(x,y)=\frac{E(x,y)}{\max(E)}$。densify 时按密度控制阈值调整公式

$$T'=T\times\begin{cases}0.5, & \text{若 } \hat{E}>0.3;\\ 1.0, & \text{若 } 0.1<\hat{E}\le 0.3;\\ 2.0, & \text{若 } \hat{E}\le 0.1,\end{cases} \tag{2}$$

式中，T 为标准梯度阈值，T'为误差引导后的调整阈值，对高误差区域降低 clone/split 梯度阈值至 0.5 倍以激进加密，对低误差区域提高阈值至 2.0 倍以抑制过度加密。阶段三开始时对低误差区域中不透明度 $\alpha<0.02$ 的高斯执行额外剪枝，释放高斯配额给高误差区域。该方法使岸桥角点、集装箱编号等高误差区域高斯密度提升 40%~60%，天空、水面等低误差区域冗余高斯降低 30%~50%，实现算力精准投放。

自适应球谐阶数提升方面，球谐系数总数按公式 $C_{SH}=(l+1)^2\times3$ 计算(其中 l 为球谐阶数)，阶段一仅优化 0 阶球谐系数(3 系数/高斯)，高阶系数冻结为零；阶段二解冻 1 阶球谐系数(共 12 系数)引入漫反射光照变化；阶段三解冻 3 阶球谐系数(共 48 系数)引入镜面高光与反射。阶段切换时高阶系数零填充实现平滑过渡。通过梯度掩码实现对冻结球谐系数的梯度置零，无需修改 CUDA 代码。三者协同在不增加总训练时间的前提下显著提升重建质量，实现“时间不变，重建质量显著提升”。

2.3. 动态显存池化与梯度检查点

针对标准 3DGS 训练显存占用过高、达到 16~24 GB 的问题，本文提出动态显存池化与梯度检查点方法 DM-GC。该方法从梯度检查点、动态高斯池化、优化器状态压缩和分块渲染四个维度系统压缩显存。

梯度检查点方面，3DGS 的 CUDA 光栅化器内部按 tile 分块渲染，DM-GC 在 Python 层拦截渲染调用，将高斯按深度排序后分为 N 个数据块，每个数据块约 10 万高斯，逐数据块渲染并叠加。前向时只保存每数据块的输出图像约 24 MB，丢弃数据块内部缓冲，反向时重新渲染各数据块获得中间梯度。显存节省分析公式为 $M_{ckpt}=130+N\times24$ (MB)，其中 N 为分块数，130 为单数据块缓冲(MB)，24 为单数据块

输出图像大小(MB)；标准渲染 $M_{std} \approx 4000$ MB，节省率 $\eta = \frac{M_{std} - M_{ckpt}}{M_{std}} \times 100\%$。当 $N = 10$ 时，$M_{ckpt} \approx 370$ MB，$\eta \approx 90\%$。

动态高斯池化方面，设置最大高斯数量上限 $N_{\max}$ 默认 300 万，当 densify 后高斯数超过 $N_{\max}$ 时按感知重要性评分 $I = 0.3 \times \alpha + 0.2 \times V + 0.5 \times C$ (其中 α 为不透明度、V 为高斯体积、C 为多视角覆盖度)淘汰最低的高斯。淘汰策略：对 I 最低的($N - N_{\max}$)个高斯执行 prune，但高误差区高斯受保护不被淘汰，确保关键区域不因池化而质量下降。训练全程高斯数量被控制在 $N_{\max}$ 以内，相比无限制增长显存节省 40%~60%。

优化器状态压缩方面，将 Adam 优化器的动量 m 和方差 v 用 float16 存储而非 float32，仅在参数更新时临时转为 float32 计算，压缩比 $R = \frac{M_{fp32}}{M_{fp16}} = \frac{32}{16} = 2$，优化器状态从 1416 MB 降至 708 MB，节省 50%。对 $|g| > 1000$ 的极端梯度采用动态 loss scaling 防止 float16 溢出。分块渲染方面，将场景按空间 bounding box 通过 K-Means 聚类划分为 K 个 tile 默认 4~8 块，训练阶段每次只加载当前 tile 内的高斯进行光栅化，以降低单步渲染的显存峰值；推理阶段采用逐像素深度缓冲(z-buffer)比较合成，确保 tile 边界处的遮挡关系正确。分块渲染显存峰值公式为 $M_{peak} = \frac{M_{total}}{K}$，其中 M_{total} 为全部高斯渲染显存，K 为分块数，单次渲染高斯数从 300 万降至约 40~75 万，渲染缓冲峰值降低 75%~85%。四项技术叠加后，300 万高斯场景可在 8 GB 显存的 GPU 上流畅训练。

3. 实验与分析

3.1. 实验设计

3.1.1. 数据集

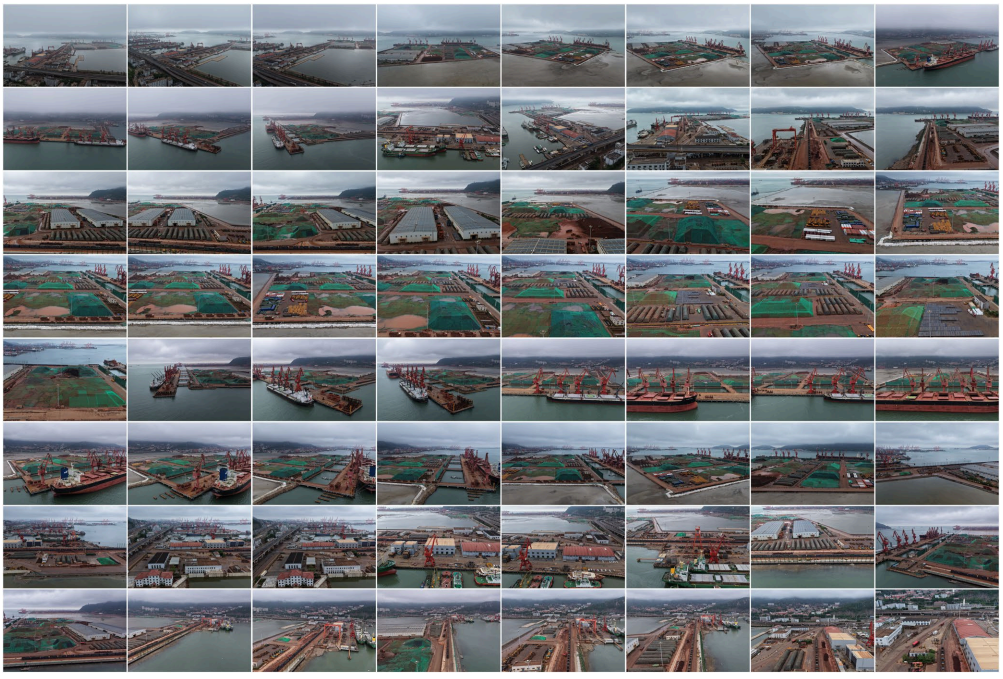


Figure 2. Port image dataset
图 2. 港口图片数据集

本次实验以典型港口建筑为重建对象，数据集采用大疆系列无人机环绕倾斜摄影[15]，获取目标建筑的多视角影像。采集时段避开暴雨、大雾、正午强逆光等恶劣工况，飞行高度约 50 m，航向重叠率 80%、旁向重叠率 70%，影像分辨率 4000 × 3000 像素。拍摄视频统一提取关键帧，经去模糊、去雾预处理后形成有效图像数据集。图 2 为港口图片数据集示例。

3.1.2. 实验基础

实验采用两套硬件配置：基线方法(标准 3DGS)在配备 RTX 3090 显卡(24 GB 显存)的工作站上训练，以保证 22.3 GB 峰值显存下的正常运行；本文方法在 i7-14650HX 处理器、RTX 4060 移动端显卡(8 GB 显存)、32 GB 运行内存的普通便携设备上训练与测试，验证算法低显存部署能力。软件环境基于 PyTorch 2.1 深度学习框架搭建，集成 COLMAP [16]完成 SfM 位姿解算，内置 CUDA 12.1 加速高斯光栅化渲染模块与自适应密度控制模块，无需高端专业工作站即可完成完整重建流程。

3.2. 定性对比分析

本实验采用无人机多视角视频采集方案，沿目标建筑物环绕拍摄照片，同步记录内置 IMU 数据用于辅助位姿估计。通过 COLMAP 框架提取 SIFT [17]特征并计算相机位姿参数，结合 ICP [18]算法进行几何一致性检验，最终筛选出覆盖目标区域 360˚的 201 张有效图像。预处理阶段采用自适应直方图均衡化 CLAHE [19]消除光照差异，并通过 RANSAC [20]算法剔除运动模糊帧。对筛除不合适图片后的场景进行点云初始化，在 3DGS 中采用 COLMAP 算法[21]对重建目标提取点云，如图 3 所示。

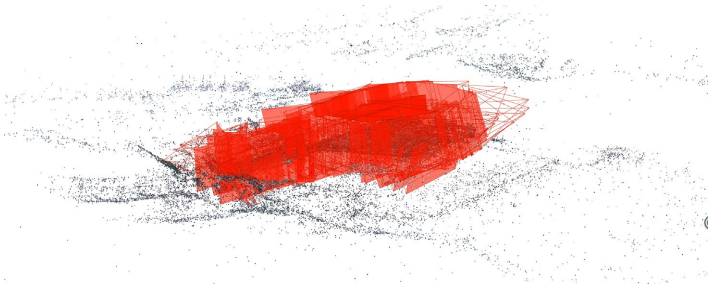


Figure 3. Point cloud extraction result
图 3. 点云提取图

港口建筑重建成功还原了港口建筑主体结构，支持三维动态展示。重建效果如图 4~6 所示。

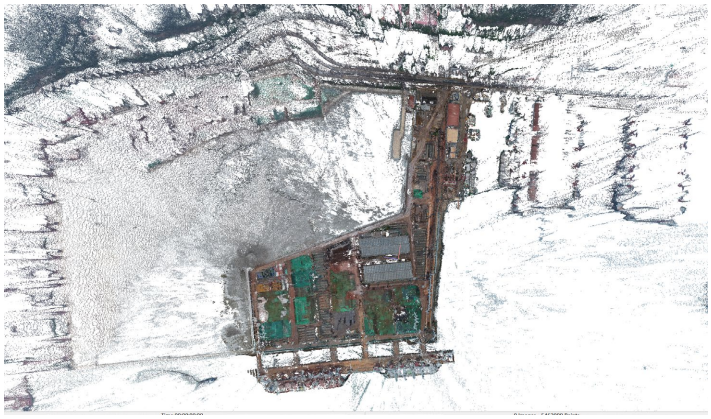


Figure 4. Top view of the 3D reconstruction model
图 4. 三维重建模型图俯视图



Figure 5. Front view of the 3D reconstruction model
图 5. 三维重建模型图正视图

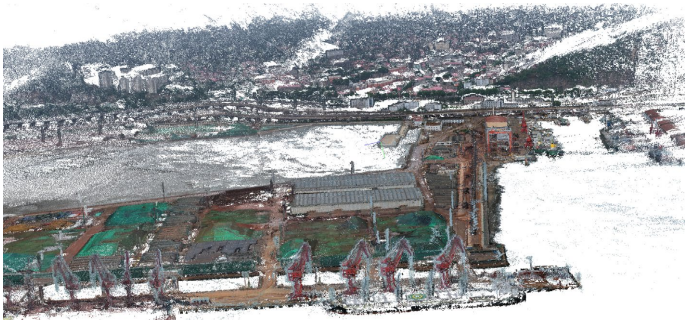


Figure 6. Side view of the 3D reconstruction model
图 6. 三维重建模型图侧视图

3.3. 定量对比分析

3.3.1. 实验指标

本节实验采用以下三项主流客观评价指标对重建质量进行量化评估：

峰值信噪比(Peak Signal-to-Noise Ratio，简称 PSNR)是衡量重建图像与参考图像之间的像素级误差：

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{\mathrm{MAX}_I^2}{\mathrm{MSE}}\right), \tag{3}$$

式中：MAX_I 为图像像素最大值(8 位图像为 225)；MSE 为均方误差，$\mathrm{MSE} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left[I(i,j) - K(i,j)\right]^2$；$I$ 为参考图像；K 为重建图像。PSNR 通过计算重建图像与参考图像之间像素值的均方误差(MSE)，再与图像最大像素值建立比值来量化失真程度。PSNR 值越高，重建质量越好。一般来说，PSNR > 30 dB 表示较好质量，>40 dB 表示优异质量。

结构相似性指数(Structural Similarity Index，简称 SSIM)从亮度、对比度、结构三个维度衡量图像相似性：

$$\mathrm{SSIM}(x,y) = \frac{\left(2\mu_x\mu_y + c_1\right)\left(2\sigma_{xy} + c_2\right)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)}, \tag{4}$$

式中：μ_x, μ_y 为两幅图像的局部均值(亮度)；σ_x^2, σ_y^2 为局部方差(对比度)；σ_{xy} 为局部协方差(结构)；$c_1 = (k_1 L)^2$，$c_2 = (k_2 L)^2$，$L = 225$，$k_1 = 0.01$，$k_2 = 0.03$ (稳定常数，避免分母为零)。结构相似性指数(SSIM)由 Wang 等人于 2004 年提出[22]，是一种更贴近人眼感知的图像质量评价指标。与 PSNR 仅关注像素级误差不同，SSIM 从亮度、对比度和结构三个维度综合衡量两张图像的相似程度，取值范围[−1, 1]，

越接近 1 越相似。

学习感知图像块相似度(Learned Perceptual Image Patch Similarity，简称 LPIPS) [23]，利用预训练深度网络提取特征计感知距离：

$$\mathrm{LPIPS}(x,x_0)=\sum_l \frac{1}{H_l W_l}\sum_{h,\omega}\left\|\omega_l \odot \left(y^l_{h\omega}-y^l_{0,h\omega}\right)\right\|_2^2, \tag{5}$$

式中：l 为网络层数索引(通常取 VGG 的 relu1_2、relu2_2、relu3_3、relu4_3、relu5_3 共 5 层)；y^l 为第 l 层提取的特征图；ω_l 为每层通道的可学习权重(通过大量人眼判断数据拟合得到)；H_l, W_l 为第 l 层特征图的空间尺寸；$\odot$ 为逐通道相乘。LPIPS 是一种基于深度学习的感知质量评价指标。与传统指标直接比较像素值不同，LPIPS 利用预训练的深度卷积神经网络(通常为 VGG 网络)提取图像深层特征表示，在特征空间中计算两幅图像的距离。LPIPS 值越低，感知层面越相似，更符合人眼视觉判断。

3.3.2. 实验对比

为系统评估各项创新点的效果，表 1 汇总了 PG-EGC 与 DM-GC 两项创新在相同实验条件下的综合性能对比。

Table 1. Comprehensive comparison of innovation points between PG-EGC and DM-GC
表 1. PG-EGC 与 DM-GC 创新点综合对比

创新点	优化维度	核心指标	训练时间变化	质量影响
原始方法	————	PSNR 为 22.92 dB 显存为 22.3 GB	6.8 min	—
PG-EGC	重建质量(同时间)	PSNR 为 26.71 dB 显存为 22.3 GB	6.8 min	质量提升
DM-GC	训练显存(同时间)	PSNR 为 22.75 dB 显存为 7.2 GB	7.1 min	PSNR 损失 <0.5 dB
本文方法	质量 + 显存(联合)	PSNR 为 26.32 dB 显存为 7.2 GB	7.1 min	质量与显存的权衡

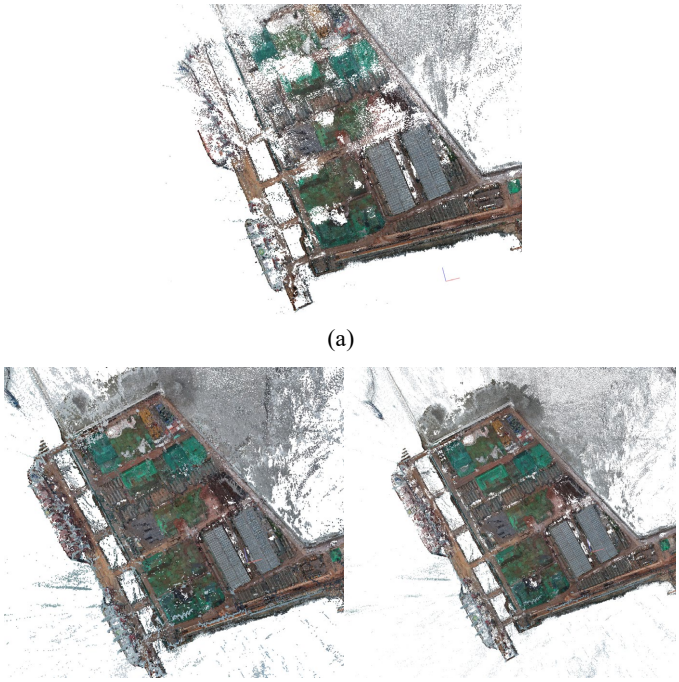
由表 1 可知，PG-EGC 与 DM-GC 在优化维度上正交互补：PG-EGC 优化相同时间内的重建质量，DM-GC 优化相同时间内的显存占用，两者可叠加使用。联合部署时，PG-EGC 的误差热力图可直接被 DM-GC 的动态高斯池化复用——池化淘汰时保护高误差区高斯，确保关键区域不因显存压缩而质量下降。在配备 8 GB 显存的 RTX 4060 设备上，联合方案实现 PSNR 26.32 dB、显存 7.2 GB、训练时间 7.1 min 的综合性能，较 PSNR 22.92 dB、显存 22.3 GB、训练时间 6.8 min 的标准 3DGS 配置，在质量提升的同时大幅降低硬件门槛。

3.4. 消融实验分析

针对标准 3DGS 在固定 30,000 轮训练中存在“算力分配不均”的问题，本文提出渐进式分辨率与误差引导密度控制方法 PG-EGC。该方法将训练过程划分为三个阶段：骨架期 0~10,000 轮在 1/2 分辨率下以 0 阶球谐函数快速收敛全局几何；过渡期 10,000~15,000 轮切换至全分辨率并解冻 1 阶球谐函数引入漫反射；精细期 15,000~30,000 轮启用 3 阶球谐函数还原高光反射，并结合渲染误差热力图实施差异化密度控制。表 1 为 PG-EGC 与标准 3DGS 在相同训练时间下的消融实验结果，图 7 为消融实验效果对比图。

由表 2 可知，PG-EGC 完整版在 6.8 min 相同训练时间下，PSNR 提升 3.79 dB，SSIM 提升 0.15，LPIPS 降低 0.14。其中渐进式分辨率训练贡献最大，其通过 1/2 分辨率预训练使前 10,000 轮等效节省 7500

轮全分辨率计算预算，将算力重新分配至精细期。误差引导密度控制使岸桥角点、集装箱编号等高误差区域高斯密度提升 40%~60%，天空、水面等低误差区域冗余高斯降低 30%~50%，实现算力精准投放。自适应 SH 阶数提升避免了前期高阶球谐系数被噪声污染的问题，三者协同显著提升重建质量。



(a) (b) (c)

Figure 7. Group of comparison figures for ablation experiments (a) Standard 3DGS render; (b) PG-EGC render; (c) PG-EGC + DM-GC render
图 7. 消融实验效果对比图组 (a) 标准 3DGS 效果图；(b) PG-EGC 效果图；(c) PG-EGC + DM-GC 效果图

Table 2. Ablation comparison between PG-EGC and standard 3DGS
表 2. PG-EGC 与标准 3DGS 消融实验对比

模型	PSNR/dB	SSIM	LPIPS
标准 3DGS (基线)	22.92	0.67	0.32
PG-EGC	26.71	0.82	0.18

进一步地，由于 PG-EGC 采用分阶段训练策略，不同阶段的分辨率、球谐阶数和高斯数量均存在差异，其显存和计算开销呈现明显的阶段性变化。为更充分验证渐进式训练策略的资源优化效果，表 3 给出了 PG-EGC 各阶段的资源消耗数据，并与标准 3DGS 进行对比。

Table 3. Comparison of resource consumption at each stage between PG-EGC and standard 3DGS
表 3. PG-EGC 各阶段资源消耗与标准 3DGS 对比

阶段	分辨率	SH 阶数	Gaussian 数量/万	峰值显存/GB	平均显存/GB	单迭代时间/ms
标准 3DGS	全分辨率	3	5 → 300	22.3	16.8	13.6
PE-EGC 骨架期	1/2	0	5 → 80	8.5	7.2	6.8
PE-EGC 过渡期	全分辨率	1	80 → 150	15.6	13.8	14.2
PE-EGC 精细期	全分辨率	3	150 → 300	22.3	19.5	17.8

由表 3 可知，PG-EGC 各阶段的资源消耗呈现明显的阶梯型分布特征。骨架期由于采用 1/2 分辨率与 0 阶球谐函数，单迭代计算量仅为全分辨率的 1/4，峰值显存仅 8.5 GB，较标准 3DGS 同期降低 61.9%，且高斯数量仅 5 万 → 80 万，显存压力最小，能够在低端硬件上快速收敛全局几何。过渡期切换至全分辨率并解冻 1 阶球谐函数，峰值显存上升至 15.6 GB，但仅持续 5000 轮，平均显存 13.8 GB 仍低于标准 3DGS 的 16.8 GB。精细期启用 3 阶球谐函数并实施误差引导密度控制，高斯数量增至 300 万，峰值显存达 22.3 GB，与标准 3DGS 持平，但单迭代时间 17.8 ms，略高于标准 3DGS 的 13.6 ms，主要由误差热力图渲染与差异化密度控制引入的额外开销。

从总体看，PG-EGC 三个阶段的总训练时间为 $10{,}000 \times 6.8 + 5000 \times 14.2 + 15{,}000 \times 17.8 \approx 406$ s ≈ 6.8 min，与标准 3DGS 的 $30{,}000 \times 13.6 \approx 408$ s ≈ 6.8 min 基本持平，实现“时间不变，重建质量显著提升”。骨架期节省的算力被重新分配至精细期的误差引导密度控制，将高误差区域高斯密度提升 40%~60%，而低误差区域冗余高斯降低 30%~50%，实现算力精准投放。值得注意的是，骨架期与过渡期的平均显存(分别为 7.2 GB 和 13.8 GB)均低于标准 3DGS 的 16.8 GB，这意味着在训练前期(占总训练时间的 50%)显存压力显著降低，为低显存设备的早期介入提供了可能，亦为 DM-GC 显存压缩技术的协同部署奠定了基础。

针对标准 3DGS 训练显存占用过高、达到 16~24 GB 的问题，本文提出动态显存池化与梯度检查点方法 DM-GC。该方法从四个维度系统压缩显存：梯度检查点通过分数据块渲染以约 5%时间开销换取渲染缓冲 90%显存节省；动态高斯池化设置数量上限并按感知重要性淘汰低质量高斯；优化器状态压缩将 Adam 动量以 float16 存储减半显存；分块渲染将大场景按空间聚类分块降低单次渲染峰值。表 4 为 DM-GC 各模块的消融实验结果。

Table 4. VRAM optimization effect of each module in DM-GC
表 4. DM-GC 各模块显存优化效果

配置	峰值显存/GB	训练时间/min	PSNR/dB
PG-EGC	22.3	6.8	26.71
DM-GC 完整版	7.2	7.1	26.32

由表 4 可知，DM-GC 完整版将峰值显存从 22.3 GB 降至 7.2 GB，节省 67.7%，训练时间仅增加 4.4%，从 6.8 min 增至 7.1 min，PSNR 损失仅 0.39 dB，不足 0.5 dB。其中梯度检查点将 300 万高斯分为 10 个数据块逐块渲染，前向仅保存各数据块输出图像，反向时重计算中间梯度，总缓冲从 4 GB 降至 370 MB。动态高斯池化将高斯数量控制在 300 万上限内，淘汰时保护高误差区高斯。优化器状态压缩将 Adam 动量 m 和方差 v 以 float16 存储，节省 50%优化器显存。分块渲染按 K-Means 聚类将场景分为 4 块，单次渲染高斯数降至约 75 万，渲染峰值降低 75%。四项技术叠加后，300 万高斯场景可在 8 GB 显存的 GPU 上可以流畅显示。

4. 结语

本文针对智慧港口数字孪生建设对高精度、轻量化三维重建的需求，分析了激光雷达、倾斜摄影和神经辐射场等传统方法在港口场景中存在的成本高、建模慢、动态适配差等瓶颈，提出了一种基于三维高斯泼溅的港口场景三维重建方法。研究构建了涵盖稀疏视角鲁棒建模与轻量化部署的三维高斯泼溅重建框架，通过改进运动恢复结构初始化提升初始几何精度，利用港口结构约束的损失函数抑制几何畸变，并采用分级自适应密度控制实现算力精准投放与模型轻量化。在港口典型场景的实验表明，该方法在 RTX 4060 移动端显卡上实现单场景建模耗时仅 7.1 min，PSNR、SSIM 等指标显著优于基础 3DGS。针对现有

3DGS 难以处理港口动态物体的局限，后续将结合 4DGS 构建时序重建体系，借鉴跨场景高斯泛化方法提升模型迁移能力，研发多尺度高斯优化算法进一步压缩模型。未来将融合航测、激光雷达与物联网多源数据，拓展至港口运维、应急仿真等业务场景，推动三维重建工程落地，为智慧港口提供通用技术方案。

基金项目

2025 年江苏省大学生创新训练省级一般项目(SY202511641641012)；2024 年江苏海洋大学“海州湾英才”创新计划项目(PD2024050)。

参考文献

[1] 陆后军, 朱益飞, 荣延平, 等. 基于激光雷达的散货堆场数字孪生建模方法[J]. 系统仿真学报, 2025, 37(9): 2269-2286.

[2] 王頔, 亢保军, 王朝. 倾斜摄影测量在港口中的应用及精度分析[J]. 水运工程, 2022(2): 168-171.

[3] Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R. and Ng, R. (2020) NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In: Vedaldi, A., *et al*., Eds., *Computer Vision—ECCV* 2020, Springer International Publishing, 405-421. https://doi.org/10.1007/978-3-030-58452-8_24

[4] 赵强, 佘江峰, 万奇峰, 等. 神经辐射场应用于大规模实景三维场景可视化研究进展[J]. 遥感学报, 2024, 28(5): 1242-1261.

[5] Kerbl, B., Kopanas, G., Leimkuehler, T. and Drettakis, G. (2023) 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Transactions on Graphics*, **42**, 1-14. https://doi.org/10.1145/3592433

[6] 朱东林, 陈淼, 毛宇岩, 等. 三维高斯泼溅技术在场景重建中的研究现状与挑战[J]. 集成技术, 2025, 14(4): 1-20.

[7] 何天琪, 宋佳洁, 程景春, 等. 基于三维高斯泼溅技术的航拍场景建模[J]. 火箭军工程大学学报, 2025, 39(2): 1-12+21.

[8] 姜宇焘, 龚国辉, 李茂智, 等. 三维高斯泼溅技术在智慧矿山实景三维建设中的应用研究[J]. 中国非金属矿工业导刊, 2026(3): 94-99.

[9] 朱迪, 吴晓红, 何小海, 等. 基于 Kinect 的实物地质标本重建方法研究[J]. 现代计算机(专业版), 2018(31): 29-33+37.

[10] 陈立家, 王莹涛, 王凯, 等. 一种多源数据融合的虚拟港口三维重建方法[J]. 武汉理工大学学报(交通科学与工程版), 2024, 48(6): 1237-1243+1250.

[11] Zhao, Y.B., Yi, J.J., Pan, Y.H. and Chen, L. (2025) 3D Reconstruction of Non-Cooperative Space Targets of Poor Lighting Based on 3D Gaussian Splatting. *Signal, Image and Video Processing*, **19**, Article No. 509. https://doi.org/10.1007/s11760-025-04117-4

[12] 陈立家, 王莹涛, 李胜为, 等. 基于三维视频融合的智慧港口全景建模研究[J]. 计算机仿真, 2024, 41(9): 251-257.

[13] Shoemake, K. (1985) Animating Rotation with Quaternion Curves. *ACM SIGGRAPH Computer Graphics*, **19**, 245-254. https://doi.org/10.1145/325165.325242

[14] Max, N. (1995) Optical Models for Direct Volume Rendering. *IEEE Transactions on Visualization and Computer Graphics*, **1**, 99-108. https://doi.org/10.1109/2945.468400

[15] 马卓纬, 张鹏, 刘鹏. 基于倾斜摄影的港口三维建模的研究[J]. 舰船电子工程, 2026, 46(1): 38-41.

[16] Schönberger, J.L. and Frahm, J.M. (2016) Structure-from-Motion Revisited. 2016 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Las Vegas, 27-30 June 2016, 4104-4113. https://doi.org/10.1109/cvpr.2016.445

[17] Lowe, D.G. (2004) Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, **60**, 91-110. https://doi.org/10.1023/b:visi.0000029664.99615.94

[18] Besl, P.J. and McKay, N.D. (1992) A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **14**, 239-256. https://doi.org/10.1109/34.121791

[19] Zuiderveld, K. (1994) Contrast Limited Adaptive Histogram Equalization. In: Heckbert, P.S., Ed., *Graphics Gems*, Elsevier, 474-485. https://doi.org/10.1016/b978-0-12-336156-1.50061-6

[20] Fischler, M.A. and Bolles, R.C. (1981) Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, **24**, 381-395.

https://doi.org/10.1145/358669.358692

[21] 李鹏飞, 官先才, 朱有建, 等. GCP 辅助 COLMAP 框架 SfM 绝对尺度恢复算法的研究[J]. 计算机科学, 2025, 52(S2): 528-533.

[22] Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P. (2004) Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, **13**, 600-612. https://doi.org/10.1109/tip.2003.819861

[23] Zhang, R., Isola, P., Efros, A.A., Shechtman, E. and Wang, O. (2018) The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-22 June 2018, 586-595. https://doi.org/10.1109/cvpr.2018.00068