

基于目标检测规避的空地协同实景三维建模优化

刘斯迪¹, 曾任之², 文泓森³, 赵梓娟⁴, 李超^{5*}

摘要: 多源数据融合、建模算法去冗余、轻量化数字传递是数字化背景下实景三维技术的新需求。基于倾斜摄影的实景三维近年发展迅速, 其建模效率提升与近地细节增强一直是实景三维领域的难点问题。针对该问题, 研究提出一种基于目标检测规避的空地协同实景三维建模低成本优化方法, 包括单目空中视角的空地协同多目扩展融合、地面视角的冗余影像的干扰目标检测筛选抽稀、空地影像融合的姿态参数预解算等技术。在保障近地建筑细节增强的情况下, 减少参与特征点建模的影像数量, 提高建模质量与效率。样例研究结果显示, 空地协同的实景三维建模经本方法抽稀并优化后, 外业地面拍摄采集的冗余工作量缩减了 37.50%, 空三计算的冗余耗时缩减了 25.00%, 模型重建的冗余耗时缩减了 35.98%, 实景三维模型冗余体积缩减了 29.69%, 优化前后实景三维模型的建筑近地细节相当。该方法的应用可为勘察设计提供一种低成本、轻量化、数字化的三维工程信息传递形式, 为新时期的城乡建成区更新设计提供高效的设计参考与测绘补充。

关键词: 空地协同; 目标检测; 实景三维; 模型优化; 数字化设计; 影像筛选

Keywords: Air-Ground Coordination; Target Detection; 3D Real Scene; Model Optimization; Digital Design; Image Selection

资助项目情况: 中建股份科技研发课题(CSCEC-2023-Z-1) 基于再生材料的新型零能耗建筑建造与无碳运行技术耦合研究与应用示范; 中建五局科技研发课题(cscec5b-2023-01) 设计建造一体化数字情境建筑多元正向 BIM 关键技术研究

引言

随着数字城市、智慧城市、AI城市的发展迭代, 城市智能模型 (City Intelligent Model, CIM) 对未来的城市数字底板提出了更高的信息化与人工智能要求^[1]。作为未来数字经济时代的时空数据基础底座, 高精度的实景三维空间信息底座建设是发展数字经济必要的基础性工作^[2]。在新时期城市更新的城乡建设背景下, 城市形象是城市全方位物态性与情境性的反映^[3], 既有立体、复杂、高动态的城市形象需要统一的实景空间信息框架以实现城市情境复现与精细化高效管理, 这对GIS、CIM、BIM等实景三维应用端技术提出了更多样的数字底板与信息更新需求。

《实景三维中国建设技术大纲 (2021版)》^[4]中定义: 实景三维 (3D Real Scene) 是对人类生产、生活和生态空间进行真实、立体、时序化反映和表达的数字虚拟空间。实景三维建模在工程图学领域

¹ 刘斯迪, 中国建筑第五工程局有限公司设计技术研究院, 工程师, Frank7Liu@163.com

² 曾任之, 中国建筑第五工程局有限公司建筑设计院, 工程师

³ 文泓森, 中国建筑第五工程局有限公司设计技术研究院, 工程师, 长沙理工大学建筑学院, 硕士研究生

⁴ 赵梓娟, 湖南农业大学东方科技学院, 讲师

⁵ 李超 (✉), 中国建筑第五工程局有限公司设计技术研究院, 高级工程师, 252429238@qq.com

应用广泛，特别是测绘领域作用巨大。近年随着无人机倾斜摄影为代表的实景三维建模技术发展，极大地便利了测绘外业的工作流程，给工程图学导入了大量实景信息。

然而，现行的主流实景三维建模技术在供给端仍存在多源数据融合、空三解算去冗、轻量化数字传递等诸多限制。(1) 多源、多视角数据融合的效率和质量需要提升。无人机、全景相机、激光扫描仪等多视角专业设备采集的数据细节多、干扰多、视角多且数据规模大、种类杂，增大了数据筛选清洗的难度和成本。(2) 数据高效建模与筛选去冗算法需要深入研究。结合在不同应用场景与采集目标，有效的压缩与去冗算法能极大压缩点云、网格、体素等常见实景三维数据模型文件体积，节省系统资源。(3) 数据分析和应用需求端的轻量化数字传递挖掘空间巨大。不限于单一展示，高效轻量的算法模型将为尺度测量、标志物定位、更新再重建等新场景应用提供有力支撑。综上，多细节、轻量化、高效率是目前实景三维优化的关键趋势。

从需求端来看，当前实景三维技术为代表的人工智能三维重建(3D Reconstruction)已在建筑学转型中赋能多类具体需求与典型应用场景，如：(1) 在设计前期勘察调研中，虽可为测绘、勘察、设计人员建立详细的场地参考，但以往单一点线面的CAD信息测绘成果交付设计已不再适应实时查看、对比与修改的转型期新设计需求；(2) 在设计成果的实景合成与效果展示中，虽合成效果逼真，但也带来了模型过大的相应算力负担；(3) 在建筑遗产保护与数字化中，虽可为建筑遗产建立详实的数字孪生体，但受限于采集视角，“飞檐”(东方)、“柱廊”(西方)等关键典型建筑遗产细节仍存在遗漏风险；

(4) 在城市片区更新中，虽实景三维可交付更新设计详细参考，但在更新设计重点关注的业态集中区，城市建筑近人尺度的昭示性牌匾、店招等“灰空间”区域，其信息缺失反而较为严重。因此，虽然当前以无人机倾斜摄影为代表的实景三维模型已交付用于转型期设计参考，但也相应面临适应设计需求新转变、模型体积轻量化、多视角协同细节增强等系列挑战。

综上，面向新时期以既有城市建筑建成区更新、再生为主的设计需求，工程图学需要一套更加便捷、高效、低成本的信息传递 workflow。同时，其实景三维的数据建模方法在近地细节增强、轻量化数字传递、建模效率提升方面应表现良好。针对此背景需求结合前人研究，以目标检测为切入点研究有干扰的冗余基础影像筛选与规避，以实现空-地协同下高效空三解算与实景三维建模生产优化。

1 相关研究进展

在总结对比目前主流实景三维建模数据源的基础上，本研究面向城市更新勘察设计需求筛选适用倾斜摄影技术，并在前人研究基础上设计空地多视角优化及地面影像筛选技术方案。

1.1 实景三维数据来源进展

目前，主流实景三维建模技术按基础数据来源与适用尺度划分，可大致分5类^[5]：

①基于卫星数据的大尺度实景三维建模(基于高分辨率卫星影像、遥感影像^[6]等数据)；②基于航空倾斜摄影的大中尺度实景三维建模(基于有人机/无人机的机载航空影像、激光雷达点云^{[8][14]}等数据)；③基于地面传感器采集的中小尺度实景三维建模(基于地面激光雷达扫描仪^{[9][10]}、移动测量车、自动驾驶汽车、智能手机、全景相机等数据)；④基于开源大数据的实景三维建模(基于 Open Street Map(OSM)^[11]、开源数字地表模型(DEM)^[12]、开源导航地图^[13]等数据)。⑤基于工程设计 BIM 模型转化的小尺度特定场景实景三维建模(基于 Revit^[14]、ArchiCAD、DP Modeler 与 3ds-Max^[22]等手工数字模型的 BIM 数据)。表 1 对比了目前主流的实景三维模型数据来源在尺度、精度、CIM 建模级别上的性能表现优劣。

表1 实景三维模型数据源优劣对比（表格来源：根据文献^[5]作者整理）

数据来源	采集途径	主要优点	主要缺点	主要适用尺度	主要建模精度	CIM级别	
天基	航天卫星数据	范围广、波段多 低成本	精度低、受遮挡 受气象干扰	全球/国域/省域	km~m	0~1	
空基	航空影像	有人机	范围广、带纹理 精度较高	高成本、高冗余 受气象干扰	区域/城市	m	2~3
		无人机	低成本、带纹理 便携性	受遮挡、高冗余 受气象干扰	片区建筑	m~dm	2~3
	航空激光	有人机	范围广、易分层 精度高、穿透性	高成本、缺纹理、 数据清洗难	区域/城市	m~dm	2~3
		无人机	低成本、易分层 精度高、穿透性	较高成本、缺纹理、 数据清洗难	片区/建筑	m~mm	3~4
地基	地面影像	低成本、带纹理 便携性	视角受限、遮挡干扰 多、高冗余	片区/建筑	m~dm	2~3	
	地面激光	精度高、易分层 速度快	高成本、缺纹理、 遮挡干扰多	片区/建筑	mm	3~4	
互联网	网络开源 大数据	低成本、范围广 数据采集便捷	精度一般、缺纹理、 数据质量差	城市/片区	m	0~1	
工程设计	BIM模型	精度高、易分层 细节多	高成本、范围小、 纹理失真	建筑	mm	4+	

注：数据参考行业主流尺度、精度、CIM级别，CIM级别参考《城市信息模型（CIM）基础平台技术导则》^[15]

1.2 实景三维建模技术进展

主流实景三维建模技术自动化为主，局部细节人工交互为辅。按照应用普及和自动化程度分5类：

①基于多源、多视角数据融合的城市建模（以倾斜摄影融合激光点云^[16]、航空影像融合DEM^[17]、开源数据融合为代表）；②基于倾斜摄影数据的点云生成技术（以光束网法平差（前方交会法）^[18]为代表）；③基于激光雷达数据的表面重建技术（以PointNet++^[19]、深度学习的神经网络点云分割法^[20]为代表）；④基于几何体块数据的纹理映射技术^[21]；⑤传统的人工手动建模技术，精度较高且效果较好，但自动化程度也最低^[22]。

以无人机倾斜摄影为主的融合建模是目前主流的片区、建筑级实景三维建模技术。随着城市化进程加快，中小尺度的片区、建筑级别的实景三维建模需求高速增长。倾斜摄影因其便携灵活、成本较低在片区建筑尺度下的实景三维建模应用广泛。同时，针对建成区更新的建设需求，无人机倾斜摄影的灵活便捷飞行采集为规划建筑从业者提供了更为灵活、低成本的数据采集方式。倾斜摄影建模成果的真实材质纹理为设计师提供了充足的实景三维信息，满足了更新设计中频繁查看、测量的实景环境的新设计需求。

虽然倾斜摄影已被广泛应用在中小尺度的城市更新设计调研中。然而，倾斜摄影受地物遮挡视角的短板却依然一直存在，成果模型的悬挑、屋檐等近地细节局部变形、拉花限制了其在城市更新中的进一步应用。而近地细节对更新设计又是必需且重要的，在规划建筑的空间设计中，近地细节往往是建筑悬挑、屋檐、骑楼、牌匾等“灰空间部位”，其空间效果也是更新设计的关注重点。因此，针对中小尺度的实景三维，基于倾斜摄影的近地细节增强是近年来学界、业界共同关注的重难点。

1.3 实景三维协同优化进展

近年，有学者尝试融合协同多源（影像+雷达）、多视角（空地协同）的采集建模方法，对倾斜摄影技术进行多视角拓展、运算去冗余优化。

(1) 影像+雷达的多源协同:

以倾斜摄影为主, 结合地面激光雷达的空地协同作业模式^{[14][16][23][24]}是目前广泛使用的协同建模方法。虽然该建模方法兼顾了模型精度与纹理细节, 且对近地细节增强效果显著, 但也存在着成本较高、外业设备不便携设站多且工作量大、内业数据清洗与融合繁琐等问题。

(2) 纯视觉领域拓展的多视角空-地协同:

对于视觉建模, 拍摄采集视角遮挡问题是影响建模成效的关键。以往通过密集采样重叠冗余覆盖不利的拍摄视角, 但也带来了空三计算与模型重建的计算负载增加, 有学者在空三解算前对拍影像组进行预筛选, 减缓了计算负载压力^[25]。而在建筑密集的城市建成区, 无人机低空飞行摄影测量的作业安全性无法得到有效保障, 建筑狭缝、檐下等灰空间区域与建筑立面、山体崖壁等垂直区域仍是飞行作业拍摄采集不利区域。

航测视角拓展: 近年测绘领域发展的数字近景摄影测量^[25]技术很好地解决了山体崖壁区、农村地区等非密集成建区的立面航测视角受限问题。但是, 在高密度的城市建成区, 线缆、行道树、道路标示牌等复杂的下垫面环境也进一步限制了贴近摄影测量的近地应用。

空地协同视角拓展: 近年在古建保护领域, 空地协同的测量方式因其檐下细节清晰、近地特征突出被广泛采用^[27]。通过地面和空中双视角的采集, 在历史建筑测绘中成效较佳, 檐下、悬挑等建筑灰空间细节信息保留完善, 实现了历史建筑的高精度、多角度的实景三维化。城市建成区的快速精细化建模亦有将此方法引入^{[28][29]}。

然而, 地面活动目标将极大影响地面拍摄采集的有效性。空地协同的倾斜摄影测量方式对地面场地拍摄要求较高, 需要封闭场地, 并搭设辅助拍摄的围护构架, 以最大程度规避环境中人、车等活动目标的干扰。因此, 对活动人车目标众多的复杂城市建成区域, 封闭场地进行测量的作业难度较大、限制多。需要一种更为自由、适用于高密度城市建成区的建筑调查方法。

研究尝试在前人研究基础上, 引入一种基于深度学习的轻量级、易部署目标检测算法, 在空地协同实景三维建模中, 基于对检出目标的规避, 对地面细节建筑影像进行去冗、抽稀。在前期外业工作中, 有效简省冗余调研工作。在后期内业工作中, 有效提升空三建模与模型重建生产的运算效率。面向新时期的更新设计需求, 以轻量级算法手段, 以期为规划建筑从业者拓展了一套链接“调研-建模-设计”环节的实景三维数字化信息传递路径。

2 方法

本方法设计优化了三部分实景三维建模技术: ①空地协同的空三解算实景三维建模, 面向实景三维近地细节增强; ②基于地面 POS 信息校正的空三参数整合, 面向协同空地影像的融合空三计算; ③基于目标检测规避的地面影像采集抽稀, 面向空地协同影像密集采样带来的建模效率提升问题。研究方法与技术框架如下图 1 所示。

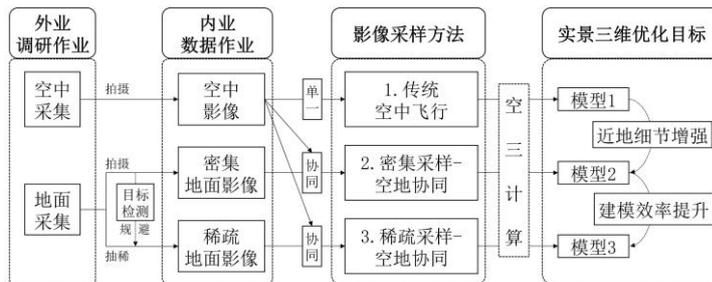


图 1 研究方法与技术框架 (图片来源: 作者自绘)

2.1 基于倾斜摄影的空三解算

光束法平差（前方交会法）是当前摄影测量和计算机视觉及机器人领域通用的一种利用影像进行定位的理论与方法^[18]，在目前主流的倾斜摄影测量中应用广泛。其基本原理如下图 2 所示，针对待测量的任意目标特征点 O ，在其至少两组拍摄采集集中， a 、 b 为对目标建筑的成像，拍摄成像点分别为 $A.(X_A, Y_A, Z_A)$ 、 $B.(X_B, Y_B, Z_B)$ 。 O 点在 a 、 b 上的成像点 O_a' 、 O_b' 即为特征点成像，当特征点成像 O_a' 、 O_b' 匹配时，空中三角形 ΔOAB 成立。

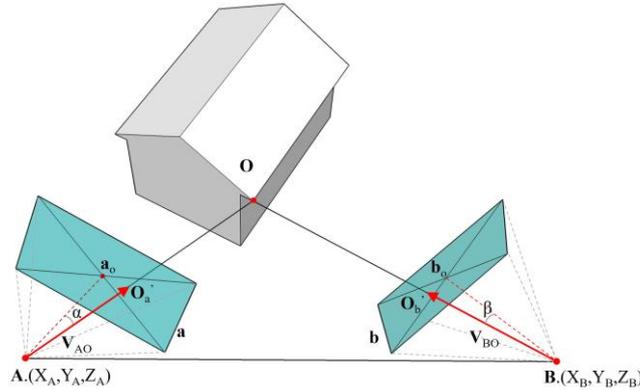


图 2 空三计算原理图（图片来源：作者自绘）

空中三角形 ΔOAB 的空间形态可借助 A 经 O_a' 至 O 的矢量 V_{AO} 与 B 经 O_b' 至 O 的矢量 V_{BO} 联合 AB 连线的矢量方向确定。 V_{AO} 、 V_{BO} 由 a 、 b 影像镜头方向和特征点 O_a' 、 O_b' 在 a 、 b 影像中的相对镜头中心点 (a_o 、 b_o) 的空间偏角 (α 、 β) 共同确定。

由于 A 、 B 位置与 AB 边长均确定，故确定形状的空中三角形 ΔOAB 的位置大小均可被解算， O 点的位置坐标 (X_O, Y_O, Z_O) 即可由以上条件联合求解。对于任意目标点 O ，拍摄成像点 AB 也是自由的，既可以是无人机拍摄的空中视角，也可以是手机、相机等拍摄的地面视角。由此，空地协同的实景三维建模的空三解算基础条件成立。

2.2 基于空地协同的空三解算

空地协同的空三建模^{[27]-[30]}采样整合了空中与地面视角优势，通过空地拍摄视角交叉，扩大了采集的目标特征点对目标建筑的覆盖。

对于传统单一无人机的空-空采样，其飞行拍摄角度、方向、重叠率均有严格要求。在对屋顶特征、立面近地特征覆盖率上表现良好。地面采集相对空中采集，由于天然的观测视角不同，特征点重叠区域能有效覆盖单一空中采集不利的建筑特征部位（屋顶、檐下、立面、地面等相互交接部位特征），是空-空采样的有效补充，如图 3。

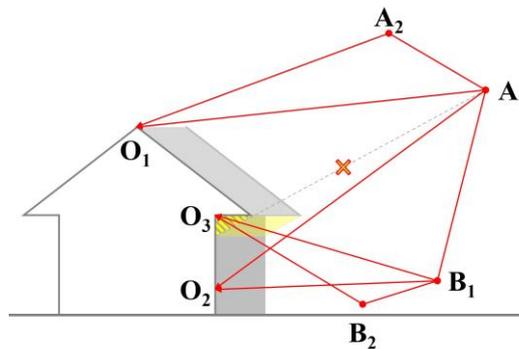


图 3 空地协同的空三计算原理（图片来源：作者自绘）

图 3 中，对于屋顶的特征点 O_1 ，基于 A_1, A_2 拍摄点的空-空采样能很好地解算屋顶的“第五立面”特征，构建三角形 $\Delta O_1A_1A_2$ 解算 O_1 定位；对于立面近地特征点 O_2 ，基于 A_1, B_1 拍摄点的空-地采样对目标建筑立面与地面的近地交接特征表现较好，构建三角形 $\Delta O_2A_1B_1$ 解算 O_2 定位；对于檐下特征点 O_3 ，其位置处于的空中 A_1, A_2 拍摄点的拍摄盲区，空-空采样的三角解算不再适用，而地面的拍摄点 B_1, B_2 恰好覆盖了这一视觉盲区，构建三角形 $\Delta O_3B_1B_2$ 用于解算 O_3 定位。

不同于传统仅有空中视角影像参与的空三解算，地面视角的加入可以有效提升建模的近地细节特征。对于不同目标特征点覆盖的建筑特征，空地协同的空三解算存在 3 种空三解算情况，如下表 2。

表 2 空-地协同的 3 种空三解算对比（图片来源：作者自绘）

目标点	O_1		O_2		O_3	
影像采样组	空-空采样		空-地采样		地-地采样	
影像点	A_1	A_2	A_1	B_1	B_1	B_2
成像位置	空中	空中	空中	地面	地面	地面
空中三角形	$\Delta O_1A_1A_2$		$\Delta O_2A_1B_1$		$\Delta O_3B_1B_2$	
特征点O的建筑特征	“第五立面”屋顶特征		立面特征、立面地面交接特征		立面、屋顶、檐下及其交接特征	

然而，空地协同虽拓展了空中拍摄的视角，地面拍摄视角受到环境中地面目标的偶发活动所频繁干扰。特别对于待更新的已建成城市建筑物密集区，密集的车流、人流等社会活动给地面拍摄带来了大量干扰目标。对目标建筑的拍摄采集易受干扰目标所阻挡，导致地面拍摄采集需要大量的冗余拍摄保障目标建筑近地细节特征。

2.3 YOLO 目标检测规避与筛选

针对地面的活动干扰目标，在拍摄采集中对干扰影像拍摄场景进行实时预筛选或在拍摄采集后对地面影像进行低效率冗余影像抽稀是在输入空三建模计算阶段前提升影像数据有效性的关键。

通过引入基于深度学习的YOLO(You Only Look Once)^[31]目标检测算法，检测拍摄场景中的无效干扰目标并规避。在空三计算前对冗余的地面信息进行筛选抽稀，降低空三运算的计算压力，保障基础影像数据的有效冗余，提升空三运算效率。

基于 YOLO 算法在地面视角活动目标的检测上性能表现优越，近年被广泛运用于自动驾驶^[32]、数字交通^{[33][34]}、工业检测^[35]等领域。YOLO 作为一种单阶目标检测器，YOLO（本研究采用 YOLOv4）由用于推论抽样特征的检测主干（Backbone）、用于扩宽感受野的层级颈部（Neck）和用于多尺度输出的预测头部（Head）等三个部分构成（图 4）。

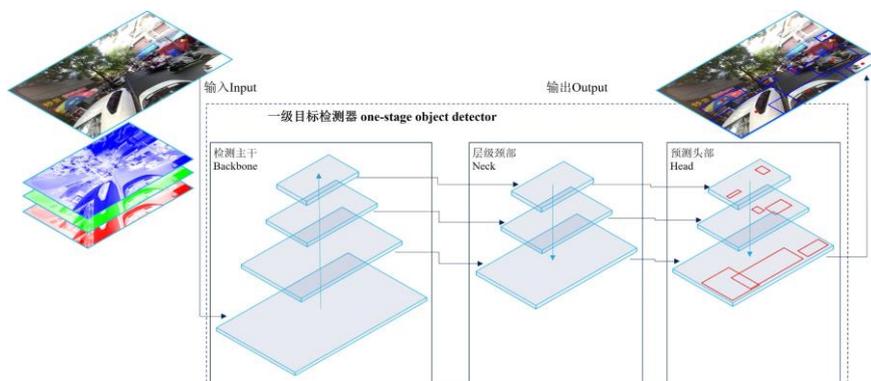


图 4 基于 YOLO 算法的目标检测框架（图片来源：作者自绘）

基于 COCO 数据集^[36]训练的 YOLO 模型，能识别包括行人、自行车、小汽车等 80 余类地物活动目标，且在性能表现上帧效率较高(YOLOv4 在 Tesla V100 GPU 上，MS COCO 数据集上的实时速度为 65 FPS，AP 为 43.5%^[31])。基于 YOLO 目标识别的结果，在前期外业务调研中可对干扰视角有效规避，

减少地面拍摄的工作量。由于 YOLO 算法的轻量化特性，在手机等消费级终端以 APP（i-Detection、YoloV53 等）形式快速部署也易于实现。

如在影像采集后的内业数据清洗中部署，能在空三解算前对外业影像数据进行有效筛选与清洗，降低空三解算与模型重建的算力消耗，提升效率。

3 实验过程与结果分析

以长沙市洋湖湿地旁某小型建筑的三维建模为例。周边道路车辆等活动目标对空地协同三维建模存在一定干扰，基于 python 语言实现 YOLO 算法对存在干扰目标的地面影像进行检测并筛选，规避有干扰的冗余地面影像，并进行空地协同实景三维建模对比实验。

3.1 实验数据

外业实验：为提高特征点的空地匹配成功率，外业空地拍摄采集均在同一时间窗口内完成，控制时间、气候、空间处于同一拍摄采集区间。空地拍摄均由低成本的消费级设备拍摄完成，空中设备为 DJI Air 2S，地面设备为 iPhone 7 Plus。

总体拍摄数量为 100 张，其中，空中影像 84 张，地面影像数量 16 张。为避免是否具备专业测绘背景带来的实验差异，空中飞手与地面拍摄者均为具备基本的建筑规划背景的从业工程师，在事前未知实验目的情况下，仅被告知需要对目标建筑及其周边环境进行拍摄，全程拍摄的空中飞行路径与地面行进路径均不作任何限制。

3.2 影像筛选与实验分组

内业实验：内业实验采用外业的 100 张利用消费级设备拍摄的影像，影像均为围绕目标建筑随机拍摄，对目标建筑所在区域进行建模，经过自动空三筛选，有 13 张影像不能参与建模，为随机拍摄干扰影像。在 16 张地面照片中，目标检测的规避算法筛选出 6 张存在活动目标的影像，剩余 10 张去冗后的地面照片。剔除的有活动目标地面影像样组及其样例如图 5(a)~(f)、6 所示。

需要说明的是，由于实验数据对照需要，本实验外业采集了全部视角的 16 张影像，在后期利用内业实验的 YOLO 检出剔除了 6 张干扰影像。而在实际工程应用中，轻量化的 YOLO 可随调研拍摄采集设备直接部署（如：YoloV53 等 APP），在外业调研中，YOLO 可实时检出有干扰目标的影像，该部分的冗余拍摄采集工作量可以提前由 YOLO 辅助规避。相应地，本实验中外业地面拍摄采集的冗余工作量可相应缩减 37.50%。

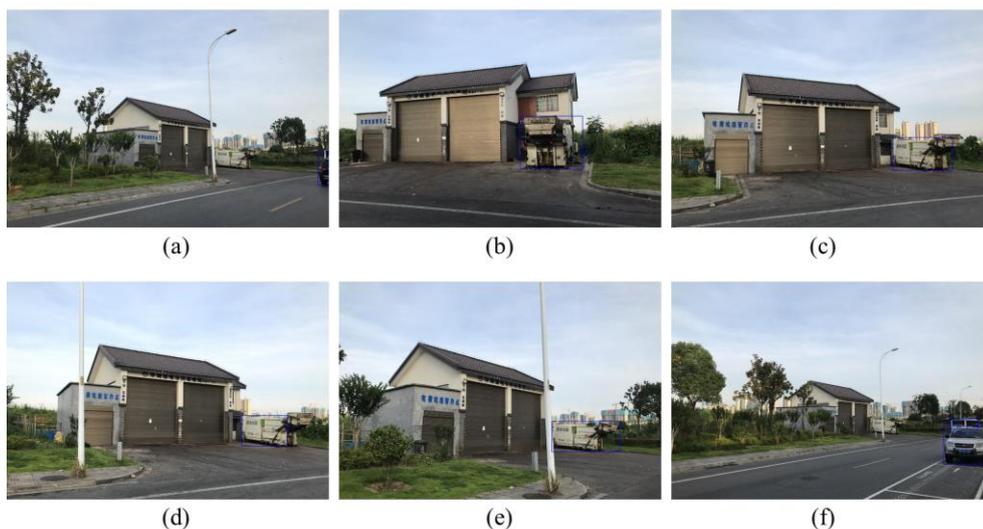


图 5 目标检测剔除的地面影像组（图片来源：作者自绘）

图 5(a)-(f)和图 6 中目标检测框(蓝色框)标识了被检出的空间环境中的活动目标,中心点(红点)表示检出目标的在该影像中视觉中心点。在筛选影像组中对场景中存在于干扰影像均予以剔除。



图 6 目标检测剔除的地面影像样例(图片来源:作者自绘)

为对比实验结果,将 84 张单一空中照片设为影像组 1,密集采样的空地协同 100(84+16)幅影像作为影像组 2,经目标检测规避抽稀的稀疏采样的空地协同 94(84+10)张影像作为影像组 3。

3.3 地面影像的相机姿态预解算

针对各组影像,由于地面影像不同于空中影像具备完全的姿态信息,地面影像仅具备位置信息(Longitude、Latitude、Height),故针对地面影像的镜头姿态旋转参数(ECEF)需要进行预先的空三解算,求解补全地面影像的 ECEF 参数。

如图 7(a)~(b),7(a)为目标检测抽稀后的地面影像组,7(b)则为抽稀前。同时,为避免其余空中影像的相机姿态参数(POS)影响建模效果,整个实验的三个影像组其余影像均使用相同的相机 POS 参数,并且在影响建模成效的各变量中固定相机位置、姿态、焦距等各参数,仅以地面影像数量作为对比结果因变量的单一初始变量。

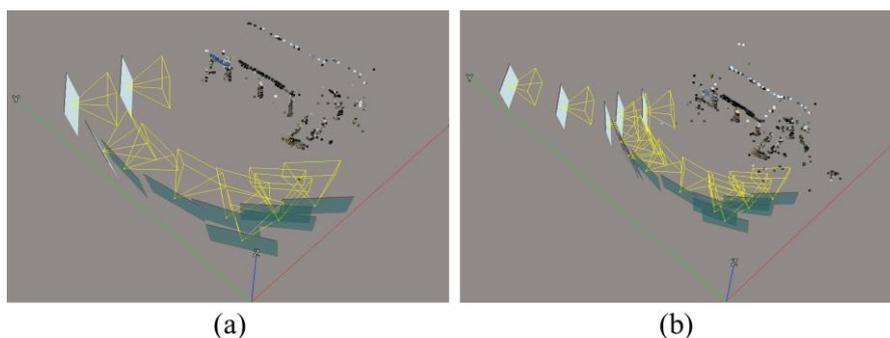


图 7 地面影像的姿态预解算(图片来源:作者自绘)

3.4 实景三维建模实验结果对比

为了更直观探究各影像组的空三建模表现,设置如下3个实验对照组:①方法1,针对单一空中飞行的影像组1进行空三建模重建;②方法2,针对密集采样影像组2,进行空地协同实景三维重建;③方法3,针对目标检测规避后的稀疏采样影像组3,进行空地协同实景三维重建。通过三组实验,

分别对比三组实验方法的实景三维建模的外调工作时长、内业建模效率、模型效果优劣。基本测试环境如下:硬件环境:CPU: Intel i9-10900k; GPU: Nvidia 2080ti; RAM: 32G。软件环境:操作系

统：某主流消费级操作系统；空三建模软件：某主流三维重建软件。所有测试硬件、软件、驱动均选用主流性能稳定版本而非最新版本，通过稳定的软硬件相互调试适配以确保实验结果均一稳定。

针对 1、2、3 影像组分别进行空中三角建模重建。对比传统空中飞行（实验组 1）、使用本方法优化的密集采样-空地协同（实验组 2）、采用本方法优化的稀疏采样-空地协同（实验组 3）。如表 3。

表 3 各影像组建模性能对比（图片来源：作者自绘）

实验编号		实验组 ①	实验组 ②	实验组 ③
影像采样方法		1.空中采样-空三解算	2.密集采样-空地协同	3.稀疏采样-空地协同
总影像数量/幅		84	100	94
地面影像数量	有干扰/幅	0	16	10
	无干扰/幅	0	10	10
空中照片数量/幅		84	84	84
地面影像空三预解算时间 Time/s		-	00:00:13/13 s	00:00:10/10 s
空三时长 Time/s	空三时长	00:01:36/96 s	00:01:48/108 s	00:01:45/105 s
	冗余耗时	0	12s	9s
	压缩/缩减率	\	100%/ 0%	75%.00/ 25.00%
无效影像数量/幅		13	13	13
自动连接点个数		28257	30413	29158
建模时长 Time/s	建模时长	00:29:05/1745 s	00:31:49/1909 s	00:30:50/1850 s
	冗余耗时	0	164s	105s
	压缩/缩减率	\	100%/ 0%	64.02%/35.98%
.Fbx 文件模型点数/vertices		750,827	847,673	822,553
.Fbx 文件模型面数/ faces		1,501,263	1,695,001	1,644,838
.Fbx 格式 模型质量效果图		参见图 8	参见图 9	参见图 10
模型 Rhino.3dm 为 例/bit	模型文件体积大小	81.2 MB:85,179,039 bit	91.7 MB:96,173,340 bit	88.6 MB:92,909,102 bit
	冗余体积	0 bit	10,994,301 bit	7,730,063 bit
	压缩/缩减率	-	100%/ 0%	70.31%/ 29.69%

注 1：为规避特征点参数缓存导致的计时误差，各组空三与重建的缓存参数均独立计算，并在独立运算结束后即实时清空参数缓存供下一组运算调用

注 2：由于 YOLO 运算的帧效率很高，遍历 16 帧的 YOLO 自身影像筛选耗时不足 300ms(不足 0.5s)，其时长不计对实验结果影响极小

在表 3 中，在利用消费级产品采样的对比研究中，各影像采样方法的性能表现如下。可以发现：传统空中采样的方法 1 建模效果一般，在地面无法采样时，能取得较为基本的测量建模成效(图 8)。实验中，将该方法作为基准对照组，相应的计算时长作为无冗余的基准值（地面空三 0s；空地协同空三 96s；建模时长 1745s）。

对于需要高精度建模，特别是需要近地细节增强的区域，密集采样-空地协同的方法 2 能在建模细节上取得良好提升(图 9)，但同时也带来了地面采集工作量、建模运算量、生成模型大小的增长。相对于传统的纯空三建模的方法 1，方法 2 质量提升的同时也增加了冗余的工作时长（地面空三冗余 13s；空地协同空三 108s，冗余耗时 12s；建模时长 1909s，冗余耗时 164s）。

而稀疏采样-空地协同的方法 3 则是兼顾了工作量增长、建模细节效果、建模效率的一种方法，在增加有限的在增加有限建模工作量的情况下，使用 YOLO 算法检测目标并规避来完成稀疏的建筑影像采样。可有效提升目标建筑的檐下、屋檐、骑楼等近地悬挑空间细节的建模质量（建模细节、精度和辨识度和点云特征点数量），并提升运算效率和缩减模型文件大小，以节省系统资源，控制空三计算与生产建模耗时，提升实景三维内业后期建模的运算效率（图 10）。相对于冗余度较高的方法 2，方法 3 压缩了该部分的冗余工作时长（地面空三由 13s 压缩至 10s；空地协同空三 105s，冗余耗时由 12s 缩减至 9s，缩减率 25%；建模时长 1850s，冗余耗时由 164s 缩减至 105s，缩减率 35.98%）。



图 8 实验组①实景三维建模效果 (图片来源: 作者自绘)



图 9 实验组②实景三维建模效果 (图片来源: 作者自绘)



图 10 实验组③实景三维建模效果 (图片来源: 作者自绘)

4 讨论

目前, 基于目标检测规避的空地协同多目视觉扩展与整合能有效增强空三计算的实景三维建模细节与效率。其主要优点有:

(1) 特征点利用率高。对于地面和空中影像的特征点利用率充分提升,特别是空地交叉区域的目标建筑细部特征解译增强。(2) 成果更新设计参考性强。成果针对城市更新设计需求更具参照性,对规划建筑从业人员关注的“人的尺度”的近地细节进行明显加强,对更新设计的参考性提升。且实景三维成果有利于后期被引入到CIM、BIM技术应用中。(3) 轻量化算法部署便携。研究采用的YOLO算法具备轻量化特点,易于借助APP部署在手机等便携终端,在外调时使用仍能实时对环境视角进行预先筛选,减轻外业地面调研工作量。(4) 算法前后端部署自由。在外调时可在拍摄时实时筛检以减少外调工作量,在外调结束后的内业数据处理中,使用本法也能有效筛选影像照片,规避无效、低效照片进入空三运算,耗费计算资源,有效提升实景三维建模效率。

作为一种低成本替代方案,本研究仍有不足:(1) 由于实验样本较小且干扰目标类型较少,在多于干扰类型的开放城市密集区建模应用中,目标检测算法的优化有效性仍需进一步的可靠性验证。同时,本方法在影像筛选条件和阈值上,按影像数量比例、干扰物视野占比、干扰物种类等筛选机制,均存在进一步研究讨论空间。(2) 由于本方法影像筛选带来的效率提升以牺牲活动干扰目标的细节特征点为代价,极限筛选抽稀减少特征点可能同时导致空三匹配失效,如何在保障空三计算匹配且准确的条件下对影像有效筛选抽稀也是潜在研究核心问题之一。(3) 随着软硬件水平的高速提升,无人机、手机为代表的空地拍摄采集设备定位精度逐步提高,YOLO为代表的目标检测技术已迭代多代(YOLOv1~v11)。受制于软硬件水平,本方法目前仅能有限提升实景三维建模的细节与效率,仍存在相当的性能优化空间,进一步的高性能软硬件适配与空地高效协作采集是未来的发展趋势。

结语

结合图像深度学习的空地协同倾斜摄影技术为实景三维建模带来了细节增强和建模效率的兼顾优化提升。

针对基于倾斜摄影的实景三维建模的近地细节与建模的效率提升的问题。研究在保障空地协同建模增强近地建筑细节的情况下,通过减少参与特征点建模的影像数量,提高建模质量与效率。在利用消费级产品采集影像的样例研究中,空地协同的实景三维建模经本方法抽稀并优化后,外业地面拍摄采集的冗余工作量缩减了 37.50%,空三计算的冗余耗时缩减了 25.00%,模型重建的冗余耗时缩减了 35.98%,实景三维模型冗余体积缩减了 29.69%,优化前后实景三维模型的近地细节相当。

研究在可能衍生的地面拍摄采集自动化、集群外业调研众包拆分、设计成果数字化传递在未来的工程图学数字化调研与实景三维建设中前景广阔。研究期望未来为勘察设计提供一种低成本、轻量化、数字化的三维工程信息获取及传递形式,架设勘察-设计数据桥梁,弥合各设计周期阶段间的数字信息鸿沟,为新时期城乡建成区更新设计提供高效的设计参考与测绘补充。

参考文献

- [1] 吴志强,甘惟,刘朝晖,等.AI城市:理论与模型架构[J].城市规划学刊,2022(05):17-23.
WU Z Q, GAN W, LIU Z H, et al. The AI city: Theory and structural model [J]. Urban Planning Forum, 2022(05):17-23(in Chinese).
- [2] 中国自然资源报,黄先锋.打造数字经济时代的时空基础底座[EB/OL].2022-10-13:[2023-04-21].
https://www.mnr.gov.cn/dt/ch/202210/t20221013_2761765.html.
China Natural Resources News, Huang X F. Building the space-time base of the digital economy era [EB/OL]. 2022-10-13:[2023-04-21]. https://www.mnr.gov.cn/dt/ch/202210/t20221013_2761765.html(in Chinese).
- [3] 杨璞.城市形象与市民社会的空间权利——中国城市设计浅析与探索[J].建筑学报,2000(09):38-41.
YANG Y. Urban image and spatial rights of civil society: A brief analysis and exploration of urban design in China[J]. Architectural Journal,2000(09):38-41(in Chinese).
- [4] 自然资源部办公厅.自然资源部办公厅关于印发《实景三维中国建设技术大纲(2021版)》的通知[EB/OL].2021-8-11:[2023-04-21].
http://gi.mnr.gov.cn/202108/t20210816_2676831.html.

- General Office of the Ministry of Natural Resources. Notice of the general office of the ministry of natural resources on printing and distributing the "3D real scene China construction technical outline (2021 edition)" [EB/OL].2021-8-11:[2023-04-21].http://gi.mnr.gov.cn/202108/t20210816_2676831.html(in Chinese).
- [5] 单杰,李志鑫,张文元.大规模三维城市建模进展[J].测绘学报,2019,48(12):1523-1541.
SHAN J, LI Z X, ZHANG W Y. Recent progress in large-scale 3D city modeling[J]. Acta Geodaetica et Cartographica Sinica, 2019, 48(12): 1523-1541(in Chinese).
- [6] Planet. Daily satellite imagery and insights[EB/OL].(2019)[2023-04-21].https://www.planet.com/.
- [7] 李德仁,刘立坤,邵振峰.集成倾斜航空摄影测量和地面移动测量技术的城市环境监测[J].武汉大学学报(信息科学版),2015,40(4):427-435,443.
LI D R, LIU L K, SHAO Z F. An integration of aerial oblique photogrammetry and mobile mapping system for urban geographical conditions monitoring [J].Geomatics and Information Science of Wuhan University,2015,40(4):427-435,443(in Chinese).
- [8] Zhang L, Li Z, Li A, et al. Large-scale urban point cloud labeling and reconstruction[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 138: 86-100.
- [9] Stoker J M, Abdullah Q A, Nayegandhi A, et al. Evaluation of single photon and Geiger mode LiDAR for the 3D elevation program[J]. Remote Sensing, 2016, 8(9): 767.
- [10] Wang Y, Chen Q, Zhu Q, et al. A survey of mobile laser scanning applications and key techniques over urban areas[J]. Remote Sensing, 2019, 11(13): 1540.
- [11] Komadina A, Mihajlović Ž. Automated 3D urban landscapes visualization using open data sources on the example of the city of zagreb[J]. KN-Journal of Cartography and Geographic Information, 2022, 72(2): 139-152.
- [12] Taillandier F. Automatic building reconstruction from cadastral maps and aerial images[J]. International Archives of Photogrammetry and Remote Sensing, 2005, 36(Part 3): W24.
- [13] 杨乃.基于空间认知的三维地图设计若干问题的研究[D].武汉大学,2010.
Yang N.Study on several issues of 3d map design based on spatial cognition[D]. Wuhan University,2010(in Chinese).
- [14] 田先斌,张永利,吴建文,等.无人机 LiDAR 场地勘测及 BIM 规划设计研究与实践[J].图学学报,2018,39(02):339-345.
TIAN X B, ZHANG Y L, WU J W, et al. Research and practice of site surveying by UAV LiDAR and BIM planning design [J]. Journal of Graphics,2018,39(02):339-345(in Chinese).
- [15] 住房和城乡建设部办公厅.住房和城乡建设部办公厅关于印发《城市信息模型(CIM)基础平台技术导则》(修订版)的通知[EB/OL].2021-6-1:[2023-4-21].https://www.mohurd.gov.cn/gongkai/zhengce/zhengcefilelib/202106/20210609_250420.html.
General Office of the Ministry of Housing and Urban-Rural Development. Notice of the general office of the ministry of housing and urban-rural development on printing and distributing the "technical guidelines for urban information modeling (cim) basic platforms" (revised edition) [EB/OL]. 2021-6-1:[2023 -4-21].https://www.mohurd.gov.cn/gongkai/zhengce/zhengcefilelib/202106/20210609_250420.html(in Chinese)
- [16] Room M H M, Anuar A. Integration of Lidar system, mobile laser scanning (MLS) and unmanned aerial vehicle system for generation of 3d building model application: A review[C]//IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2022, 1064(1): 012042.
- [17] 李德仁.论地球空间信息的3维可视化:基于图形还是基于影像[J].测绘学报,2010,39(02):111-114.
LI D R.3D visualization of geospatial information:graphics based or imagery based [J].Acta Geodaetica et Cartographica Sinica,2010,39(02):111-114(in Chinese).
- [18] 单杰.光束法平差简史与概要[J].武汉大学学报(信息科学版),2018,43(12):1797-1810.
SHAN J.A brief history and essentials of bundle adjustment[J]. Geomatics and Information Science of Wuhan University,2018,43(12):1797-1810(in Chinese).
- [19] Qi C R, Yi L, Su H, et al. Pointnet++: Deep hierarchical feature learning on point sets in a metric space[J]. Advances in neural information processing systems, 2017, 30.
- [20] Hu Q, Yang B, Khalid S, et al. Towards semantic segmentation of urban-scale 3D point clouds: A dataset, benchmarks and challenges[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 4977-4987.
- [21] Weinhaus F M, Devarajan V. Texture mapping 3D models of real-world scenes[J]. ACM Computing Surveys (CSUR), 1997, 29(4): 325-365.
- [22] 张涛,相诗尧,李振江,等.面向公路施工现场的三维实景模型建模方法[J].公路,2020,65(10):224-229.
ZHANG T, XIANG S Y, LI Z J, et al. Modeling method of 3D real scene for construction site of highway[J].Highway,2020,65(10):224-229(in Chinese).
- [23] 谢云鹏,吕可晶.多源数据融合的城市三维实景建模[J].重庆大学学报,2022,45(04):143-154.
XIE Y P,LYU K J. City 3D real-world modeling based on multi-source data fusion [J].Journal of Chongqing University,2022,45(04):143-154(in Chinese).
- [24] 高利敏,李俊杰,李文清,等.无人机倾斜摄影结合激光扫描仪三维逆向建模[J].测绘通报,2021(02):161-163.
GAO L M, LI J J, LI W Q, et al.3D reverse modeling of unmanned aerial vehicle tilt photography with laser scanner [J].Bulletin of surveying and mapping,2021(02):161-163(in Chinese).
- [25] 张明磊,张云生,邹峥嵘.一种面向倾斜影像三维建模的影像筛选方法[J].测绘工程,2018,27(03):68-71+80.
ZHANG M L, ZHANG Y S, ZOU Z R. An image selection method for 3D modeling of oblique images[J]. Engineering of Surveying and Mapping,2018,27(03):68-71+80(in Chinese).

- [26] 张祖勋.数字摄影测量的发展与展望[J].地理信息世界,2004(03):1-5.
ZHANG Z X. Future development and prospect of digital photogrammetry [J].GEOMATICS WORLD,2004(03):1-5(in Chinese).
- [27] 孙保燕,杨正阳,陈款,等.融合航摄影像与地面照片三维重建技术在考古中的应用[J]. 科学技术与工程,2019,19(17):262-266.
SUN B Y, YANG Z Y, CHEN K, et al. Application of three-dimensional reconstruction technology combining aerial photography and ground photography in archaeology[J]. Science Technology and Engineering,2019,19(17):262-266(in Chinese).
- [28] 张强,钟季廷,吴少平,等.一种基于空地融合的精细化实景三维建模方法:中国,CN111540048A[P].2020-08-14.
Zhang Q, Zhong J T, Wu S P, et al. A refined 3D real scene modeling method based on space-ground fusion: China, CN111540048A[P]. 2020-08-14(in Chinese).
- [29] 李闻杰,左小清,唐灵,等.手机影像辅助倾斜摄影的快速精细化建模方法研究[J].城市勘测,2022(03):60-64.
Li W J, Zuo X Q, Tang L, et al. Research on fast fine modeling method of mobile image assisted tilt photography [J].Urban Geotechnical Investigation Surveying,2022(03):60-64(in Chinese).
- [30] 唐志军,王海峰,张畅,等.基于手机影像实景精细化建模局部快速更新的方法:中国,CN113963047A [P].2022-01-21.
Tang Z J, Wang H F, Zhang C, et al. A method for local rapid update of real-scene refined modeling based on mobile phone images: China, CN113963047A [P]. 2022-01-21(in Chinese).
- [31] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.
- [32] 邹斌,林思阳,尹智帅.基于 YOLOv3 和视觉 SLAM 的语义地图构建[J].激光与光电子学进展,2020,57(20):124-130.
ZOU B, LIN S Y, YIN Z S. Semantic mapping based on YOLOv3 and visual SLAM[J]. Laser & Optoelectronics Progress,2020,57(20):124-130(in Chinese).
- [33] 赵璐璐,王学营,张翼,张美月.基于 YOLOv5s 融合 SENet 的车辆目标检测技术研究[J].图学学报,2022,43(05):776-782.
ZHAO L L, WANG X Y, ZHANG Y, et al. Vehicle object detection based on YOLOv5s fusion SENet [J].Journal of Graphics,2022,43(05):776-782(in Chinese).
- [34] 杨大慎,陈科圻,马翠霞.一种面向无人机视频的多尺度摘要的设计与实现[J].图学学报,2020,41(02):224-232.
YANG D S, CHEN K Q, MA C X. Design and implementation of a multi-scale summarization for unmanned aerial vehicle videos[J].Journal of Graphics,2020,41(02):224-232(in Chinese).
- [35] 李建明,杨挺,王惠栋.基于深度学习的工业自动化包装缺陷检测方法[J].包装工程,2020,41(07):175-184.
LI J M, YANG T, WANG H D. An industrial automation packaging defect detection method based on deep learning [J].PACKAGING ENGINEERING,2020,41(07):175-184(in Chinese).
- [36] Lin T Y, Maire M, Belongie S, et al. Microsoft coco: Common objects in context[C]//Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13. Springer International Publishing, 2014: 740-755.

英文摘要

Air-Ground Coordination 3D Real Scene Modeling Optimization Based on Object Detection and Avoidance

Abstract: Multi-source data fusion, modeling algorithm de-redundancy, and lightweight digital transmission are new requirements for 3D real scene technology under the digital background. 3D real scene based on oblique photography has developed rapidly in recent years, and its modeling efficiency improvement and near-ground detail enhancement have always been difficult problems in the field of 3D real scene. Aiming at this problem, the study proposes a low-cost optimization method for air-ground collaborative 3D real scene modeling based on target detection and avoidance, including air-ground collaborative multi-eye extended fusion of monocular aerial perspective, and interference target detection screening and thinning of redundant images from ground perspective. Attitude parameter pre-calculation of air-ground image fusion and other technologies. This method reduces the number of images involved in feature point modeling while enhancing the near-ground details of buildings and improves the quality and efficiency of modeling. The results of the sample study show that air-ground coordinated 3D real scene modeling was extracted and optimized by this method, the redundant workload of external ground photography acquisition is reduced by 37.50%, the redundancy time of aerial triangulation is reduced by 25.00%, the redundancy time of model reconstruction is reduced by 35.98%, the redundant volume of the 3D real scene model is reduced by 29.69%, the near-ground details of the buildings in the 3D real scene model before and after optimization are comparable. The application of this method can provide a low-cost, lightweight, digital 3D engineering information transmission form for survey and design, and provide an efficient design reference and surveying and mapping supplement for the renewal design of urban and rural built-up areas in the new era.

相关专利

本文技术已申请公开相关专利, 如需引用请参考: [1]杨瑛,刘斯迪,胡月明,等.一种基于目标检测规避的空地协同实景三维建模优化方法[P].湖南省:CN202310065153.3,2023-04-28.