

通信行业 天面类资源识别与三维重建

Identification and 3D Reconstruction of Antenna Platform Resources in Communication Industry

张昀玮, 竹梦圆(中讯邮电咨询设计院有限公司, 北京 100048)

Zhang Yunwei, Zhu Mengyuan(China Information Technology Designing & Consulting Institute Co., Ltd., Beijing 100048, China)

摘要:

为了实现对现存无线网资源更好的数字化管理,通常需要借助三维重建技术^[1]完成其数字孪生。但是在天面类资源的应用场景下,由于其种类的多样性以及环境的复杂性,简单采用无人机或全景相机进行3D建模的方法并不能普遍应用。随着人工智能技术的兴起,借助参数识别而非扫描的方法可以更加精准且迅速的实现天面类资源的三维重建。经过数据采集与实地实验,结合识别算法与预置模型组装可以很好的完成通信行业天面类资源的三维重建任务。

关键词:

人工智能; 三维重建; 深度估计; 目标检测

doi: 10.12045/j.issn.1007-3043.2023.07.007

文章编号: 1007-3043(2023)07-0037-06

中图分类号: TN914

文献标识码: A

开放科学(资源服务)标识码(OSID):



Abstract:

In order to achieve better digital management of existing wireless network resources, it is usually necessary to complete its digital twin with the help of 3D reconstruction technology. However, in the application scenario of antenna platform resources, because of the diversity of types and the complexity of the environment, the method of simply using panoramic cameras for 3D modeling cannot be widely applied. With the rise of artificial intelligence technology, the 3D reconstruction of antenna platform resources can be realized more accurately and quickly with the help of parameter identification instead of scanning. After data collection and field experiments, the recognition algorithm combined with component assembly can well complete the 3D reconstruction task of antenna platform resources in the communication industry.

Keywords:

Artificial intelligence; 3D reconstruction; Depth estimation; Object detection

引用格式: 张昀玮, 竹梦圆. 通信行业天面类资源识别与三维重建[J]. 邮电设计技术, 2023(7): 37-42.

0 引言

目前通信行业的资源可视化手段主要为数据各类形式展现和GIS^[2]呈现2种方式,还未实现实物世界与数字世界的一一对应。随着VR、3D和传感技术的发展,虚拟数字世界与真实世界的互动成为可能^[3]。但是在前期调研工作中发现,天面类资源因其所在室外环境的不确定性与复杂性,目前3D建模主要采用无人机^[4],但因无人机操控复杂、影像3D无法实现自动化处理以及图像识别等关键技术数据匮乏,未能满足

普及推广使用的条件。

本文借助时下热门的人工智能算法,通过将天面类资源参数拆分,探索出一套天面类资源自动化3D建模方案,可以实现基站铁塔和天线的数字孪生,推动网络资源可视化管理。

1 天面类资源三维建模研究发展现状

目前业界室外场景的三维建模方法主要分为2种:一种为将场景内容显式地表达为点云、网格、体素等元素的集合^[5-6],另一种则借助神经网络将场景隐式地抽象为一个连续的5D神经辐射场(Neural Radiance Field, NeRF)^[7-8],通过大量参数来拟合场景内光线的

收稿日期: 2023-05-18

传播情况。在显式建模中,关键技术主要分为2步,分别为运动恢复结构(Structure from Motion, SfM)^[9-10]的稀疏重建与多视图立体(Multi View Stereo, MVS)的稠密重建。在稀疏重建中,关键的技术点是需要通过匹配不同图像中的同一关键点^[12-13]来还原不同图片相机的位姿信息。但是在室外天面类资源中,因为其表面纹理特征不清晰,外加大量的仰拍天空占比,从图像中还原位姿变得困难无比,参考实地拍摄图片(见图1),图像内存在大量的树木与天空噪声,难以匹配关键点信息。而另一种NeRF算法,由于需要多角度的隐式模式匹配,其输入往往需要进行多角度的环绕拍摄图像,而在室外对近40 m高的铁塔进行环绕拍摄对于拍摄硬件与人体条件的要求又过分苛刻,此方法也同样不能用于室外天面类资源的三维重建场景中。



图1 真实室外天面类资源场景

2 天面类资源三维建模新方案

在天面类资源的数字孪生过程中,工作人员并不会细致地关注铁塔上的走线或杂物情况,其资源关注点主要为铁塔种类、天线挂高以及铁塔每层的挂杆占用率。这些使得天面类资源参数化成为可能,即在资源的数字孪生体中主要表现出塔的种类、平台层数、天线挂高以及占用情况即可,无需做到细致如塔顶鸟巢的建模。因此在多次实地实验与算法验证后,本文

提出一种全新的天面类资源三维建模方案:通过将铁塔拆分为高度、类别2种维度的信息,分别训练高度估计与目标检测2种人工智能算法,实现铁塔的关键信息识别;在信息提取完全后,再利用这些矢量信息配合标准模型进行三维矢量建模。本方案无需人员操纵复杂的机器如无人机,也无需提供多角度复杂的环绕拍摄,只需要采集人员携带智能手机等拍摄设备,进行整体塔身的拍摄以及平台细节的拍摄,即可完成整个天面类资源的参数建模。整个建模流程如图2所示。

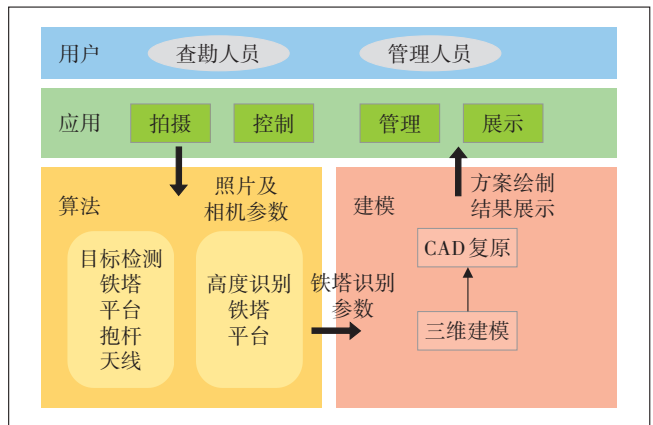


图2 方案流程

2.1 高度估计算法

本文提出了一套无需标定的、基于单摄像头的铁塔高度估计算法,此算法主要包括目标检测、语义分割、深度估计、三维坐标还原和高度估计5个模块。目标检测模块可以根据图像实现图像中不同铁塔个体像素位置的推理,其中像素坐标系即为每个像素点在图像中的位置 (u, v) ,其位置定义如图3所示;语义分割模块则可以实现单目RGB图像中推理镂空铁塔的掩膜;深度估计模块则可以根据单目RGB图像还原出图像中各点像素间的深度关系;三维坐标还原模块可以根据智能手机的焦距、陀螺仪参数生成投影矩阵,进一步借助深度估计模块生成的深度关系还原出图像中各点像素的高度关系;最后,高度估计模块可根据已知的参照物确定准确的像素间高度映射,实现待检测物体最高像素与最低像素间的高度差计算。经过以上几个模块间的推理与信息利用,本方案最终可以推理出待检测物体的真实高度,系统整体流程如图4所示。

在目标检测模块中,本方案利用yolov8目标检测算法,在室外铁塔数据集上进行训练。该数据集来自

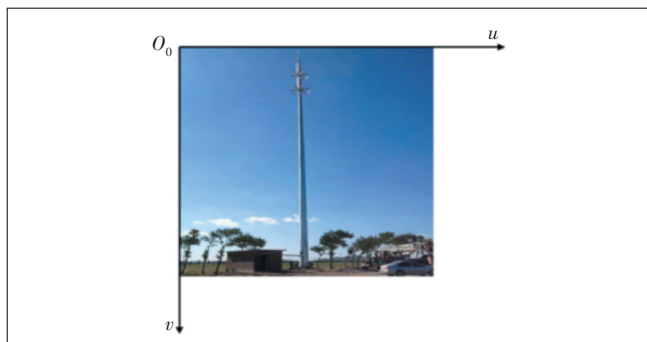


图3 图像像素坐标系

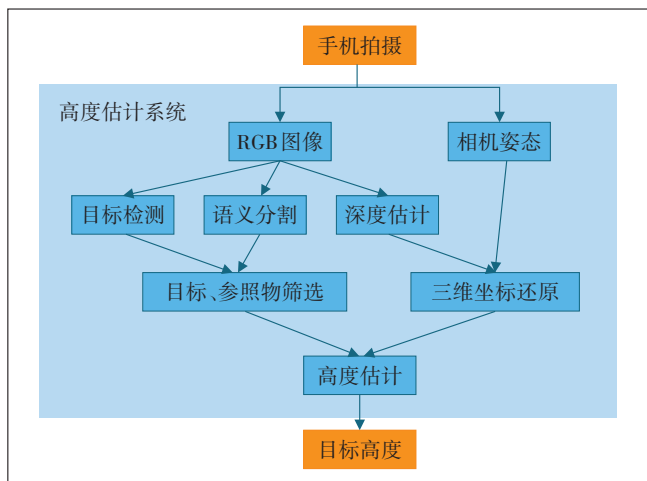


图4 铁塔高度估计系统

无线传输等业务部门,数据集分为塔身主体图像以及平台细节图像2种,其中整体图像被用于铁塔高度估计算法。在本模块,通过对输入的铁塔整体图片进行特征提取与解码,可输出图像中不同种类的铁塔及其外接矩形4个角点的像素坐标。基于此,本方案可以获得图像中不同铁塔个体的种类及其像素位置,为后续三维坐标还原提供位置和分类信息。图5为目标检测模块演示。

在语义分割模块中,本方案采用 HRNet48+OCR



图5 目标检测模块演示

的语义分割算法,在自有的铁塔语义分割数据集上进行训练,最终实现对输入图像中铁塔形状的抠出。在实际应用场景中,由于部分铁塔的钢架结构,不可避免地会造成图像中镂空的出现,如果不进行语义分割,则极易引入天空部分的噪声,造成高度估计精度下降。图6为语义分割模块演示。



图6 语义分割模块演示

在深度估计模块中,本方案采用 boostingMonocularDepth 算法进行深度关系估计。该算法可仅依靠单张图像作为输入,实现图像中每个像素的深度估计。值得一提的是,本算法输出深度数值并无量纲,深度信息经过归一化后可以作为关系信息加以利用。图7为深度估计模块演示。



图7 深度估计模块演示

在三维坐标还原模块中,本系统采用了针孔成像模型,像素坐标系到世界坐标系的转换关系如式(1)所示。本方案利用铁塔作为垂直于地面的钢体结构当作先验信息,并假设世界坐标系 z 轴与大地 z 轴平行,坐标系原点与相机坐标系原点重合,依此构建了相机坐标系与世界坐标系间的转换关系,其旋转平移矩阵的旋转参数通常可由智能手机的陀螺仪传感器读取,平移系数为0。根据具体的姿态角可以生成垂直于地面的世界坐标系到相机坐标系的转换矩阵。相机坐标系中像平面与真实物平面关系如图8所示,

根据相机坐标系到像素坐标系的转换关系〔见式(2)],在已知像素坐标系中 u, v, C_u, C_v 的情况下,只需知道 f_x, f_y 的比例关系,就可知道相机坐标系下目标的高度坐标的相对信息。深度信息 Z 并非准确的具体信息,因而可以根据传感器信息,直接读取 f_x, f_y 比例关系进行坐标的等比例缩放。最后可根据像素坐标系下的像素坐标以及深度关系,利用内参矩阵和旋转矩阵的逆求出各个像素对应的世界坐标系下的位置关系信息。

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (1)$$

$$\begin{cases} \frac{X}{Z} \times f_x + C_u = u \\ \frac{Y}{Z} \times f_y + C_v = v \end{cases} \quad (2)$$

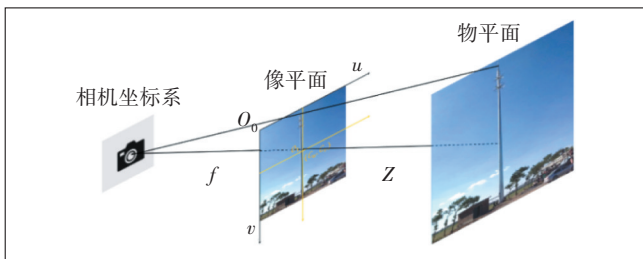


图8 相机坐标系

最后,在高度估计模块中,利用先前各个模块的输出结果与转换公式,本方案可以轻松求得图像中各个像素点间归一化后的高度差值,只需要提供图像中已知高度的参照物信息,便可借由缩放系数获得铁塔真实高度的估计。

2.2 目标检测算法

为了能够实现天线、抱杆等组件参数的识别,本方案还囊括了具体组件的细节识别算法实现。为了实现铁塔组件细节的识别,本方案利用前文提到的实地拍摄图像构建了细节检测数据集。数据集按照图像的涵盖内容分为2个部分,第1部分为上文提到的铁塔塔身整体数据集,包含了单管塔(有平台)、独杆塔(无平台)、三管塔、角钢塔、路灯杆、景观塔、仿生树、方舱等几种标准塔形与平台层和非平台层这2种平台类型。利用整体数据集可以获得图像中不同铁塔的种类以及铁塔上的平台分层情况,即使无平台也可根据“非平台层”确认直接固定在塔身的抱杆分层。

考虑到远景拍摄抱杆、天线等组件极为不清晰的情况,笔者根据手机数字变焦后拍摄的细节图像创建了第2部分的细节数据集,囊括了如天线、RRU、AAU、抱杆、美化罩、平台、小平台等组件,用于实现具体组件的细节识别。

在细节识别部分,本文仍沿用整体检测的思路,使用yolov8算法作为检测器进行训练。yolov8算法借助其全新的Decoupled-Head提高了检测头部的精度,并且创新性地使用VFL Loss作为类别损失,使用DFL Loss + CIUO Loss作为位置损失,有效地提供了一种解决正负样例不均衡的思路,在本方案中可以很好地解决图像中RRU、抱杆等小目标的检出问题。

2.3 三维重建算法

如上文所述,本方案能够成功实施全部依赖于构建的铁塔抽象参数集。将需要总结的参数总结成表,并根据经验与标准制作相关的三维模型,其中塔身高度模块根据语义分割结果与高度估计结果可输出塔身最高点与最低点之间的差值,并结合经验模型获得最终塔身高度;抱杆与平台分布以及天线与抱杆分布则可根据图像中的位置关系进行匹配。最终建模效果是标准模型与识别参数结果的共同呈现。

获取参数识别结果之后,本方案根据标准模型构建了不同组件三维模型,如不同高度的塔身与避雷针模型,除此之外还构建了如不同直径的平台与不同长度的抱杆模型,最终结合识别参数与经验进行组件间的组装。安装顺序为根据高度优先构建塔身与避雷针,而后向上组装不同高度的平台与抱杆,最后向对应抱杆上组装天线、AAU等设备。表1为铁塔参数抽象示意。

3 算法实验结果

本方案基于自有的室外铁塔数据集进行相关算法的训练。数据集共为2种三类算法进行赋能训练,本文在整体塔身图像的组件标注与塔身抠图上训练了铁塔整体的检测算法与语义分割算法,在铁塔细节图像的组件标注上训练了铁塔细节的检测算法,每部分模型均取得了较高的准确率。

3.1 整体塔身语义分割算法

将训练集、测试集以9:1的比例进行划分,划分完之后训练集约630张、测试集约70张。为使模型在测试集上的结果具有参考性,设定测试集完全对模型不可见,严格保密。在此数据集上,进行不同 backbone

表1 铁塔参数抽象示意

组件类型	识别参数	经验模型
塔身种类	三管塔、路灯杆等	塔身标准模型
塔身高度	高度数值, 如 36.5 m	离散型数值, 如 35 m、40 m
避雷针长度	高度数值	离散型数值, 如 3 m、5 m
抱杆长度、支架横撑	无	根据塔身种类与平台情况决定
平台直径	无	根据塔身种类与平台情况决定
平台层、无平台层高度	高度数值, 根据位置推算	情况较复杂
平台层、无平台层抱杆数量	抱杆对象数量	有标准模型, 如 3、6 等
设备类型	组件对象检测结果	根据标准模型建模
抱杆空置情况	是否有天线、RRU 等	无经验模型

与 head 的实验, 表 2 与表 3 证明了选取的 HRNet48+OCR 的结果优越性。最后, 经过不同的数据增强与损失函数优化, 模型性能满足使用需求。

表2 不同 backbone 实验结果

backbone	head	epoch	miou
ResNet50	UperNet	46	38.19
HRNet48	UperNet	36	39.50
ConvNeXt-base	UperNet	32	37.19
Swin Transformer-base	UperNet	30	38.23

表3 不同 head 实验结果

backbone	head	epoch	miou
HRNet48	UperNet	36	39.5
HRNet48	DeepLabv3+	38	35.9
HRNet48	Unet	32	32.5
HRNet48	OCRnet	46	43.9

由于铁塔塔身在整体图像的分布情况较为集中, 且像素数量远远小于背景像素数量, 因此借助 lovasz loss 来解决正负例间分布不均衡的问题, 最终分割效果如图 9 所示, 其中半透明蓝色掩膜代表铁塔塔身, 半透明绿色掩膜代表背景。最终实验结果如表 4 所示。

3.2 整体与细节检测算法

将图片按照所拍主体分为整体图像与细节图像, 并对其进行了不同类别的标注, 以构建数据集。在训练过程中, 对 2 种检测模型采用了同样的算法选型, 并都参照 7.5:2.5 的比例划分了训练集与测试集, 其中整体目标检测算法测试指标如表 5 所示。

细节图像主要包含了铁塔上部的功能组件标注, 为了还原抱杆与平台层的位置关系, 将平台层同样作



图9 分割实验结果

表4 最终实验结果

backbone	head	loss	数据增强	图片大小 (batchsize)	miou
HRNet48 (加载预训练权重)	OCRnet	0.1 CE loss+1.0 lovasz loss	随机缩放、多尺度训练、随机反转、随机旋转、随机裁剪、颜色对比度的变化	512x1 024 (16)	81.98

表5 整体目标检测算法指标

种类	实例数量	Bbox 准确率	回归率
三管塔	5	0.816	0.895
角钢塔	2	0.986	0.857
单管塔(有平台)	15	0.996	1.000
独杆塔(无平台)	8	0.963	0.750
路灯杆	8	0.990	0.875
景观塔	8	0.812	1.000
仿生树	1	0.821	1.000
方舱	2	0.886	1.000
平台层	44	0.949	0.839
无平台层	99	0.950	0.770
汇总	197	0.917	0.899

为对象进行了标注, 最终细目标检测算法指标如表 6 所示, 检测结果如图 10 所示。

表6 细目标检测算法指标

种类	实例数量	Bbox 准确率	回归率
天线	1 649	0.909	0.867
抱杆	2 236	0.828	0.882
平台	248	0.973	0.968
RRU	1 172	0.790	0.619
美化罩	43	1.000	0.848
AAU	2 651	0.807	0.850
汇总	7 999	0.885	0.796



图 10 检测结果

3.3 最终建模效果

在多个塔站中进行图像收集与最终建模测试,实验表明,本方案可以较为直接地展示塔站资源的分布和闲置情况,工作人员可以较为便捷地自动化采集铁塔数据,构建天面类资源三维模型,实现天面类资源的数字孪生。最终三维建模结果如图 11 所示。



图 11 最终三维建模结果

4 结语

随着通信行业的发展,一套能够在复杂场景下实现自动化3D采集建模的设备亟需研发落地,以应对行业内资产统计与发展的新需求。在AI行业突飞猛进的今日,数字孪生的概念越发流行。本方案基于时下

新兴的人工智能算法,结合多种场景下的不同任务,在真实业务数据集上训练天面类资源识别模型,并配合经验与标准模型,构建了一套高效而便捷的通信行业天面类资源重建算法方案。该方案可以快速帮助公司与个人在无任何外加条件的情况下创建与控制资产的数字孪生体,为未来完全化的数字孪生奠定可能。

参考文献:

- [1] 熊思博,王琦,刘光洁. 三维重建技术的发展与现状研究综述[J]. 电脑知识与技术,2022,18(36):114-117.
- [2] 孙传博. 基于GIS技术的通信行业基础资源管理系统[D]. 哈尔滨:哈尔滨理工大学,2019.
- [3] 周晓雨,林娴,李志扬. VR中真实场景的多视图三维重建[J]. 科技与创新,2019(6):114-116.
- [4] 于广瑞,刘兴春,时春霖,等. 一种稳健的无人机影像三维重建方法[J]. 测绘通报,2022(6):76-81.
- [5] CHARLES R Q, SU H, KAICHUN M, et al. PointNet: deep learning on point sets for 3D classification and segmentation [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017:77-85.
- [6] QI C R, YI L, SU H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, California, USA: Curran Associates Inc., 2017: 5105-5114.
- [7] MILDENHALL B, SRINIVASAN P P, TANCIK M, et al. NeRF: representing scenes as neural radiance fields for view synthesis [J]. Communications of the ACM, 2022, 65(1):99-106.
- [8] ZHANG K, RIEGLER G, SNAVELY N, et al. Nerf++: analyzing and improving neural radiance fields [DB/OL]. [2023-04-02]. <https://arxiv.org/abs/2010.07492>.
- [9] ULLMAN S. The interpretation of structure from motion [J]. Proceedings of the Royal Society of London. Series B. Biological Sciences, 1979, 203(1153):405-426.
- [10] SCHÖNBERGER J L, FRAHM J M. Structure-from-motion revisited [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016:4104-4113.

作者简介:

张昀玮,助理工程师,硕士,主要从事人工智能计算机视觉算法工作;竹梦圆,工程师,硕士,主要从事计算机视觉、视频AI领域的算法研发工作。

