

卫星单目位姿估计的关键点检测与不确定度同步预测

苏昂^{1,2}, 王梓^{1,2*}, 王靖皓^{1,2}, 李璋^{1,2}

(1. 国防科技大学 空天科学学院, 湖南 长沙 410073; 2. 图像测量与视觉导航湖南省重点实验室, 湖南 长沙 410073)

摘要: 卫星单目视觉位姿估计通常先提取卫星图像关键特征点, 再求解 n 点透视 (perspective- n -points, PnP) 问题得到摄像机与卫星之间的相对位姿关系, 其中卫星关键点检测的精度是决定单目位姿估计精度的关键。针对该问题提出一种卫星关键点高精度检测方法, 该算法在预测关键点图像坐标的同时给出关键点坐标预测的不确定度, 在此基础上构建加权的 PnP 约束方程求解相对位置和姿态, 显著提升了卫星单目位姿估计精度。在公开的卫星单目位姿估计 SPEED 数据集上开展了相关实验, 实验结果表明提出的同步预测关键点坐标与不确定度的卫星关键点检测方法能够有效提升关键点检测精度, 并且通过求解加权的单目位姿估计问题显著提升了卫星单目位姿估计精度。

关键词: 关键点检测; 不确定度预测; Transformer; 单目视觉; 卫星位姿估计

中图分类号: V248.1 文献标志码: A 文章编号: 1001-2486(2025)02-098-11

论文
拓展

Keypoints detection and uncertainty synchronous prediction for satellite monocular pose estimation

SU Ang^{1,2}, WANG Zi^{1,2*}, WANG Jinghao^{1,2}, LI Zhang^{1,2}

(1. College of Aerospace Science and Engineering, National University of Defense Technology, Changsha 410073, China;

2. Hunan Key Laboratory of Videometrics and Vision Navigation, Changsha 410073, China)

Abstract: Satellite monocular pose estimation usually extracts the keypoints of the satellite in the images, and then solves the PnP (perspective- n -points) problem to obtain the relative position and attitude between the camera and the satellite, in which the accuracy of satellite keypoints detection is the key to determine the accuracy of monocular pose estimation. To solve this problem, a high-precision satellite keypoints detection method was proposed, which predicted the image coordinates of the keypoints and gave the uncertainty of keypoints coordinate prediction synchronously. Then, a weighted PnP constraint equation was constructed to solve the relative position and attitude on this basis, which significantly improved the accuracy of satellite monocular pose estimation. Relevant experiments were carried out on the public satellite monocular pose estimation dataset named SPEED. The experimental results show that the proposed keypoints detection method for synchronously predicting keypoints coordinates and their uncertainty can significantly improve the accuracy of satellite keypoints detection, and by solving the weighted monocular pose estimation problem, the accuracy of satellite monocular pose estimation has been significantly improved.

Keywords: keypoints detection; uncertainty prediction; Transformer; monocular vision; satellite pose estimation

卫星单目视觉位姿估计研究使用单目摄像机拍摄卫星并估计摄像机相对于卫星的三维位置和姿态, 在交会对接、编队飞行和空间目标抓捕等领域都有着重要应用。视觉位姿估计方法具有高精度、低功耗和体积质量小等优点, 非常适合太空环境应用需求。其中单目视觉测量方法只需一个摄像头, 与双目立体视觉相比, 进一步降低了设备复

杂度和安装空间需求。

基于单目视觉的位姿估计一般通过求解 n 点透视 (perspective- n -points, PnP) 问题^[1-3] 解决, 假设卫星外观特征关键点在其本体坐标系下的三维空间坐标已知, 在图像上提取卫星关键点二维图像坐标, 构建 PnP 约束方程并求解, 可以得到摄像机与卫星之间的相对位置和姿态

收稿日期: 2024-01-10

基金项目: 国家自然科学基金资助项目 (12302252)

第一作者: 苏昂 (1987—), 男, 河南周口人, 副教授, 博士, E-mail: suang@nudt.edu.cn

*通信作者: 王梓 (1994—), 男, 山东菏泽人, 助理研究员, 博士, E-mail: wangzi16@nudt.edu.cn

引用格式: 苏昂, 王梓, 王靖皓, 等. 卫星单目位姿估计的关键点检测与不确定度同步预测[J]. 国防科技大学学报, 2025, 47(2): 98-108.

Citation: SU A, WANG Z, WANG J H, et al. Keypoints detection and uncertainty synchronous prediction for satellite monocular pose estimation[J]. Journal of National University of Defense Technology, 2025, 47(2): 98-108.

等 6 自由度位姿参数。对于卫星单目位姿估计任务,在图像中如何实现高可靠和高精度的卫星特征关键点检测成为决定卫星单目位姿估计精度的关键。

对于合作卫星,可以在卫星表面放置人工设计的合作标志辅助实现高精度的关键点检测,比如对角^[4]、圆形^[5]或科尔多瓦大学增强现实库(augmented reality library from the University of Cordoba, ArUco)标志^[6]等,在交会对接等任务中通常采用这种方式。然而,空间中多数卫星是非合作的,其表面不方便或无法放置合作标志,需要利用卫星外观的纹理或结构特征。传统方法主要采用人工设计特征提取和匹配卫星特征关键点,常用的特征有:尺度不变特征^[7](scale invariant feature transform, SIFT)、加速分割测试特征^[8](features from accelerated segment test, FAST)、二进制鲁棒独立初等特征^[9](binary robust independent elementary features, BRIEF)、旋转 FAST 和 BRIEF^[10](oriented FAST and rotated BRIEF, ORB)、二进制鲁棒不变可扩展关键点^[11](binary robust invariant scalable keypoints, BRISK)等。在空间任务中,卫星关键点自动提取面临着卫星结构复杂、视角和亮度变化、强光干扰以及复杂地球背景等困难,传统人工设计特征在性能上存在一定不足。

随着深度学习技术的发展,在图像分类、目标检测和跟踪等计算机视觉任务中,神经网络的方法展示出了强大性能。随后产生了一系列基于深度学习的单目视觉位姿估计方法^[12-13],根据是否显式求解 PnP 问题,这些方法可分为直接法和间接法。直接法使用神经网络直接预测出卫星相对位姿,比如:文献[14]采用高斯混合模型和软标签实现姿态直接预测;文献[15]采用分类和回归两种方式直接预测相对姿态,再采用高斯牛顿法迭代求解相对位置。间接法先使用神经网络预测卫星关键点坐标,再求解 PnP 问题得到卫星相对位姿,比如:文献[16-18]采用热图(Heatmap)实现关键点检测;文献[19-21]采用坐标回归的方式直接输出关键点坐标;文献[22]先回归出关键点坐标,再采用位姿回归网络得到位姿。

深度卷积神经网络(convolutional neural network, CNN)处理图像时,卷积核覆盖范围有限,对长距离依赖关系建模能力较弱。针对该问题,文献[23]借鉴自然语言处理领域对语句单词序列建模的 Transformer 模型,提出了端到端目标检测框架 DETR(detection transformer)。在此基

础上产生了基于 Transformer 模型的卫星单目位姿估计方法^[24-25],采用 Transformer 模型直接预测卫星关键点坐标,再求解 PnP 问题得到卫星相对位姿参数。

以上基于深度学习的单目位姿估计间接方法使用预测关键点求解卫星位姿时,忽视了关键点坐标预测精度的差异性,求解 PnP 问题时无法根据关键点坐标预测精度对各关键点构建的约束方程进行加权,影响了卫星单目位姿估计精度。针对这个问题,以基于 Transformer 模型的卫星单目位姿估计方法^[24]为基础,本文提出一种高精度的关键点检测算法,在预测关键点坐标的同时给出关键点坐标预测的不确定度,然后构建加权的 PnP 约束方程求解卫星相对位姿参数。对于提出的方法,使用欧空局和斯坦福大学联合发布的航天器位姿估计数据集^[26](spacecraft pose estimation dataset, SPEED)开展了充分的实验验证,实验结果表明该方法能显著提升关键点检测精度和卫星单目位姿估计精度。

1 卫星关键点高精度检测

1.1 基于 Transformer 的关键点检测框架

采用基于 Transformer 模型的卫星关键点检测框架,与文献[24]中的方法类似,模型总体框架如图 1 所示,主要由主干网络、编码器、解码器和预测网络等部分构成。

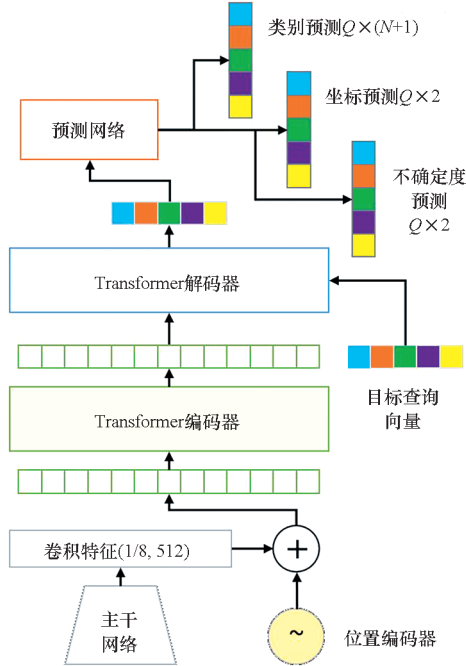


图 1 基于 Transformer 的卫星关键点检测框架
Fig. 1 Transformer-based satellite keypoints detection framework

主干网络用来从输入图像中提取卷积特征,针对关键点检测任务,既要保证提取特征具有足够空间分辨率,又不能太大导致模型训练占用显存过高。以 ResNet-50 为基础设计的主干网络只使用原网络的 Layer2 和 Layer3。原网络 Layer2 和 Layer3 的特征图尺寸分别是输入图像的 1/8 和 1/16,将 Layer2 通过 1×1 卷积调整特征通道数为 256,将 Layer3 通过 2 倍上采样和 3×3 卷积得到 1/8 特征图并调整特征通道数为 256,再将 Layer2 和 Layer3 两层输出的特征合并,最终输出的卷积特征图尺寸为输入图像的 1/8,特征通道数为 512。

主干网络提取的卷积特征与空间位置编码输入 Transformer 编码器得到编码特征向量,编码特征向量与目标查询向量输入 Transformer 解码器得到解码特征向量,解码特征向量输入预测网络得到卫星关键点类别预测、坐标预测以及坐标不确定度预测结果。假设模型中目标查询向量数量为 Q ,表示最终输出 Q 个关键点检测结果,卫星关键点的数量为 N ,则类别预测分支、坐标预测分支和坐标不确定度预测分支的特征维度分别为 $Q \times (N+1)$ 、 $Q \times 2$ 和 $Q \times 2$ 。

1.2 卫星关键点表示方法

卫星关键点表示包括类别项、坐标项和坐标不确定度项等。类别项表示为独热编码形式,设卫星上有 N 个关键点,令 i 为关键点类别索引,增加背景类别,则 i 的取值范围为 1 到 $N+1$ 。记第 i 个关键点独热编码为 $\mathbf{c}_i \in \mathbb{R}^{N+1}$,如式(1)所示。

$$\mathbf{c}_i(k) = \begin{cases} 1 & k=i \\ 0 & k \neq i \end{cases} \quad (1)$$

卫星关键点表示的坐标项即为关键点二维图像坐标,记为 $\mathbf{p}_i = [x_i, y_i]$ 。卫星关键点坐标不确定度项为预测的关键点二维图像坐标标准差,记为 $\mathbf{d}_i = [\sigma_x^i, \sigma_y^i]$ 。那么卫星关键点可表示为类别项 $\mathbf{c}_i$ 、坐标项 $\mathbf{p}_i$ 和坐标不确定度项 $\mathbf{d}_i$ 的组合。

每幅图像中,卫星关键点训练数据真值是包括 N 个元素的关键点集合,每个元素均由类别项、坐标项组成。记该集合为 Φ ,如式(2)所示。

$$\Phi = \{t_i | t_i = (\mathbf{c}_i, \mathbf{p}_i)\}_{i=1}^N \quad (2)$$

式中: t_i 表示第 i 个关键点元素; i 的取值范围为 1 到 N ; $\mathbf{c}_i$ 是第 i 个关键点的独热编码,表示其属于卫星每个关键点(含背景类别)的概率; $\mathbf{p}_i$ 表示第 i 个关键点的二维图像坐标。

基于 Transformer 的卫星关键点检测模型输出的预测结果是包含 Q 个元素的关键点集合,每个元素均由类别项、坐标项和坐标不确定度项组

成。记该集合为 Ω ,如式(3)所示。

$$\Omega = \{h_j | h_j = (\mathbf{c}_j, \mathbf{p}_j, \mathbf{d}_j)\}_{j=1}^Q \quad (3)$$

式中: h_j 表示预测的第 j 个关键点元素; j 的取值范围为 1 到 Q ; $\mathbf{c}_j$ 表示第 j 个预测属于卫星各个关键点的概率(含背景类别); $\mathbf{p}_j$ 和 $\mathbf{d}_j$ 分别表示第 j 个预测关键点的图像坐标和坐标不确定度。

1.3 关键点坐标与不确定度同步预测

计算关键点检测模型训练的损失函数时,先将关键点真值集合 Φ 和预测集合 Ω 各元素进行匹配对应,匹配对应方法以及类别预测损失函数与 DETR 方法^[23]和文献[24]中的方法一致。为描述关键点坐标预测不确定度,将关键点二维图像坐标描述为两个独立的一维高斯分布,这种表示方法在目标检测中已有成功应用^[27]。以关键点图像坐标 x 分量为例,高斯分布函数 $N(x_e, \sigma^2)$ 定义如式(4)所示。

$$P_G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x_e)^2}{2\sigma^2}} \quad (4)$$

式中:高斯分布均值 x_e 为关键点坐标预测值;标准差 σ 表示关键点坐标预测不确定度。标准差越小,意味着坐标预测越准确。对于训练数据中的关键点坐标真值,可描述为 $\sigma \rightarrow 0$ 时的高斯分布,此时即为狄拉克 δ 函数(单位冲击函数),定义如下:

$$P_D(x) = \delta(x - x_g) \quad (5)$$

式中: x_g 为关键点坐标真值。

计算关键点坐标真值与预测值两个概率分布之间的 KL 散度(Kullback-Leibler divergence)作为坐标预测损失。忽略 KL 散度中无关项,可得坐标预测损失函数为:

$$L_{\text{loc}} = \frac{(x_g - x_e)^2}{2\sigma^2} + \frac{1}{2} \ln(\sigma^2) \quad (6)$$

式(6)所示的损失函数由两项组成,其中第二项控制坐标预测标准差尽量小,第一项自适应调节坐标预测值与预测标准差。当预测值 x_e 与真值 x_g 相差较大时,预测标准差 σ 倾向于取一个较大的值以使损失函数减小;当预测值 x_e 与真值 x_g 相差较小时,预测标准差 σ 倾向于取一个较小的值。值得注意的是,当 σ 等于 1 时,式(6)所示的损失函数将退化为常用的 2 范数损失函数。

如图 2 所示,绿色竖直箭头表示狄拉克 δ 函数,代表关键点位置真值,两条曲线(实曲线和虚曲线)表示模型预测关键点坐标与不确定度的高斯函数。采用式(6)所示损失函数训练优化过程中,当预测值与真值相差较大时模型预测结果倾向于为图中的实曲线,当预测值与真值相差较小

时模型预测结果倾向于为图中的虚曲线。

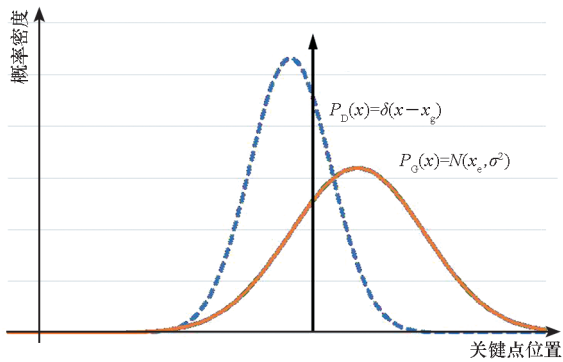


图2 关键点坐标与不确定度预测损失函数示意图

Fig.2 Diagram of the loss function for keypoints coordinates and uncertainty prediction

将式(6)所示损失函数分别对关键点坐标预测值 x_e 和关键点坐标预测标准差 σ 求偏导,结果如下:

$$\frac{\partial L_{loc}}{\partial x_e} = \frac{x_g - x_e}{\sigma^2} \quad (7)$$

$$\frac{\partial L_{loc}}{\partial \sigma} = -\frac{(x_g - x_e)^2}{\sigma^3} + \frac{1}{\sigma} \quad (8)$$

可以发现,由于关键点坐标预测标准差 σ 出现在分母中,在模型刚开始训练时可能产生梯度爆炸现象而导致模型训练时不稳定。因此,在模型训练时,可令 $\alpha = \ln(\sigma^2)$,让网络输出 α 而不是直接输出 σ 以避免梯度爆炸现象,此时损失函数如式(9)所示。

$$L_{loc} = \frac{e^{-\alpha}}{2} (x_g - x_e)^2 + \frac{1}{2} \alpha \quad (9)$$

采用与 smooth L_1 类似的策略,当 $|x_g - x_e|$ 大于某个阈值 β 时将 2 范数损失改为 1 范数,最终模型训练的坐标预测损失函数为:

$$L_{loc} = \begin{cases} e^{-\alpha} \frac{1}{2\beta} (x_g - x_e)^2 + \frac{\alpha}{2} & |x_g - x_e| < \beta \\ e^{-\alpha} \left(|x_g - x_e| - \frac{\beta}{2} \right) + \frac{\alpha}{2} & |x_g - x_e| \geq \beta \end{cases} \quad (10)$$

2 卫星加权单目位姿估计

2.1 单目位姿估计问题

卫星单目位姿估计根据卫星关键点空间三维坐标到图像二维坐标的成像关系建立约束方程来求解得到摄像机与卫星之间的相对位姿,通常称为 PnP 问题。假设卫星本体坐标系到摄像机坐标系的姿态旋转矩阵为 $\mathbf{R} \in \mathbb{R}^{3 \times 3}$, 卫星本体坐标系原点在摄像机坐标系中的位置向量(也叫作平移向量)为 $\mathbf{T} \in \mathbb{R}^{3 \times 1}$, 分别记为:

$$\mathbf{R} = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \quad (11)$$

$$\mathbf{T} = [t_x \quad t_y \quad t_z]^T \quad (12)$$

设卫星第 i 个关键点在其本体坐标系下的空间三维坐标为 $\mathbf{P}_i = [X_i, Y_i, Z_i]^T$, 该关键点对应的图像二维坐标为 $\mathbf{p}_i = [x_i, y_i]^T$, 则卫星关键点的成像模型可用式(13)所示的共线方程表示。

$$\begin{cases} x_i = f_x(\mathbf{P}_i) = F_x \frac{r_1 X_i + r_2 Y_i + r_3 Z_i + t_x}{r_7 X_i + r_8 Y_i + r_9 Z_i + t_z} + C_x \\ y_i = f_y(\mathbf{P}_i) = F_y \frac{r_4 X_i + r_5 Y_i + r_6 Z_i + t_y}{r_7 X_i + r_8 Y_i + r_9 Z_i + t_z} + C_y \end{cases} \quad (13)$$

式中: C_x 和 C_y 为摄像机像主点; F_x 和 F_y 为摄像机等效焦距; $f_x(\cdot)$ 和 $f_y(\cdot)$ 分别表示图像坐标系 x 和 y 方向的成像函数。当已知摄像机像主点和等效焦距等内参数、卫星关键点在卫星本体坐标系的空间三维坐标以及卫星关键点对应的图像二维坐标时,可通过求解式(14)所示最优化问题得到卫星本体坐标系与摄像机坐标系之间的相对位姿关系 $\mathbf{R}$ 和 $\mathbf{T}$ 。

$$\arg \min_{\mathbf{R}, \mathbf{T}} \sum_{i=1}^N \left\| \begin{bmatrix} x_i - f_x(\mathbf{P}_i) \\ y_i - f_y(\mathbf{P}_i) \end{bmatrix} \right\|_2 \quad (14)$$

2.2 加权单目位姿估计问题

如果卫星关键点检测方法无法获取关键点图像坐标预测的不确定度,比如文献[24]的方法,则求解单目位姿估计问题的各关键点权重均相同,如式(14)所示。本文提出的卫星关键点高精度检测方法在预测关键点坐标的同时能够给出关键点坐标预测不确定度,表示为坐标预测结果的标准差 $\mathbf{d}_i = [\sigma_x^i, \sigma_y^i]$ 。那么,式(14)所示的最优化问题中,可以根据各关键点坐标预测的不确定度设置约束方程权重,再求解加权的单目位姿估计问题得到更高精度的位姿测量结果。加权的单目位姿估计最优化问题表示如下:

$$\arg \min_{\mathbf{R}, \mathbf{T}} \sum_{i=1}^N \left\| \begin{bmatrix} \lambda_x^i (x_i - f_x(\mathbf{P}_i)) \\ \lambda_y^i (y_i - f_y(\mathbf{P}_i)) \end{bmatrix} \right\|_2 \quad (15)$$

式中: λ_x^i 和 λ_y^i 分别表示第 i 个关键点 x 和 y 方向的权重。

2.3 加权单目位姿估计问题求解

对于式(14)和式(15)所示的非线性最小二乘优化问题,通常将其一阶泰勒展开转化为线性最小二乘问题,然后迭代求解。设摄像机与卫星

之间 6 自由度的相对位置和姿态参数构成待求解向量为 $\mathbf{X} \in \mathbb{R}^{6 \times 1}$, 其中姿态参数的表达形式有多种, 包括欧拉角、罗德里格斯和四元数等。求解单目位姿估计问题时采用罗德里格斯法, 罗德里格斯旋转向量为三维列向量, 其中向量的模表示旋转角度, 向量的方向表示旋转轴, 该方法将卫星与摄像机之间的姿态旋转关系描述为绕空间某个旋转轴旋转一定角度。

$f_x(\mathbf{P}_i)$ 和 $f_y(\mathbf{P}_i)$ 对待求解位姿参数 $\mathbf{X}$ 求偏导的雅可比向量分别为 $\mathbf{A}_x^i \in \mathbb{R}^{1 \times 6}$ 和 $\mathbf{A}_y^i \in \mathbb{R}^{1 \times 6}$, 则式(14)一阶泰勒展开转化的线性最小二乘问题如下所示:

$$\begin{cases} x_i = x_i^0 + \mathbf{A}_x^i(\Delta\mathbf{X}) + n_x^i \\ y_i = y_i^0 + \mathbf{A}_y^i(\Delta\mathbf{X}) + n_y^i \end{cases} \quad (16)$$

式中, x_i^0 和 y_i^0 分别是位姿初值条件下关键点在图像上重投影 x 和 y 方向坐标, $\Delta\mathbf{X}$ 为待求解位姿相对于位姿初值的偏差, n_x^i 和 n_y^i 分别是关键点 x 和 y 方向图像坐标测量误差。考虑全部 N 个卫星关键点, 式(16)所示线性最小二乘问题写成矩阵形式为:

$$\begin{bmatrix} x_1 - x_1^0 \\ y_1 - y_1^0 \\ \vdots \\ x_i - x_i^0 \\ y_i - y_i^0 \\ \vdots \\ x_N - x_N^0 \\ y_N - y_N^0 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_x^1 \\ \mathbf{A}_y^1 \\ \vdots \\ \mathbf{A}_x^i \\ \mathbf{A}_y^i \\ \vdots \\ \mathbf{A}_x^N \\ \mathbf{A}_y^N \end{bmatrix} \Delta\mathbf{X} + \begin{bmatrix} n_x^1 \\ n_y^1 \\ \vdots \\ n_x^i \\ n_y^i \\ \vdots \\ n_x^N \\ n_y^N \end{bmatrix} \quad (17)$$

式(17)记为 $\mathbf{Y} = \mathbf{A}\Delta\mathbf{X} + \mathbf{B}$, 其中 $\mathbf{Y} \in \mathbb{R}^{2N \times 1}$ 为测量向量, $\mathbf{A} \in \mathbb{R}^{2N \times 6}$ 为观测矩阵, $\mathbf{B} \in \mathbb{R}^{2N \times 1}$ 为测量误差向量。根据设计关键点坐标及其不确定度预测损失函数时的假设, 各关键点坐标测量误差 n_x^i 和 n_y^i 分别满足零均值方差为 $(\sigma_x^i)^2$ 和 $(\sigma_y^i)^2$ 的独立高斯分布, 其中标准差 σ_x^i 和 σ_y^i 即为本文提出算法中定义的关键点坐标预测不确定度。测量误差 $\mathbf{B}$ 的协方差矩阵 $\mathbf{W} = \text{Cov}(\mathbf{B}) = E(\mathbf{B}\mathbf{B}^T)$ 即为如式(18)所示的对角阵:

$$\mathbf{W} = \text{diag}[(\sigma_x^1)^2, (\sigma_y^1)^2, \dots, (\sigma_x^i)^2, (\sigma_y^i)^2, \dots, (\sigma_x^N)^2, (\sigma_y^N)^2] \quad (18)$$

设 $\mathbf{D} = [\sigma_x^1 \ \sigma_y^1 \ \cdots \ \sigma_x^i \ \sigma_y^i \ \cdots \ \sigma_x^N \ \sigma_y^N]^T$, 则式(18)可简写为 $\mathbf{W} = \mathbf{D}\mathbf{D}^T$ 。在式(17)左右两边同时乘以 $\mathbf{D}^{-1}$, 将转化为如下所示的最小二乘问题:

$$\mathbf{D}^{-1}\mathbf{Y} = \mathbf{D}^{-1}\mathbf{A}\Delta\mathbf{X} + \mathbf{D}^{-1}\mathbf{B} \quad (19)$$

令 $\mathbf{Y}^* = \mathbf{D}^{-1}\mathbf{Y}$, $\mathbf{A}^* = \mathbf{D}^{-1}\mathbf{A}$, $\mathbf{B}^* = \mathbf{D}^{-1}\mathbf{B}$, 则式(19)

可简写为:

$$\mathbf{Y}^* = \mathbf{A}^* \Delta\mathbf{X} + \mathbf{B}^* \quad (20)$$

对于式(20)所示的最小二乘问题, 其测量误差 $\mathbf{B}^*$ 的协方差矩阵 $\mathbf{W}^* = \text{Cov}(\mathbf{B}^*)$ 为:

$$\mathbf{W}^* = \mathbf{D}^{-1}\mathbf{W}(\mathbf{D}^{-1})^T = \mathbf{D}^{-1}\mathbf{D}\mathbf{D}^T(\mathbf{D}^{-1})^T = \mathbf{I} \quad (21)$$

此时测量误差的协方差矩阵为单位对角阵。根据最小二乘理论, 可得式(20)所示线性最小二乘问题的无偏最优估计为:

$$\Delta\mathbf{X} = ((\mathbf{A}^*)^T \mathbf{A}^*)^{-1} (\mathbf{A}^*)^T \mathbf{Y}^* \quad (22)$$

因此, 根据以上最小二乘理论, 式(17)两边同时乘的矩阵 $\mathbf{D}^{-1} = [1/\sigma_x^1 \ 1/\sigma_y^1 \ \cdots \ 1/\sigma_x^i \ 1/\sigma_y^i \ \cdots \ 1/\sigma_x^N \ 1/\sigma_y^N]^T$ 即为该最小二乘问题的权重。该权重为关键点坐标预测标准差的倒数, 进一步将权重归一化, 可得式(15)加权单目位姿估计最优问题中的权重为:

$$\begin{cases} \lambda_x^i = \frac{1/\sigma_x^i}{\sum_{i=1}^N 1/\sigma_x^i} \\ \lambda_y^i = \frac{1/\sigma_y^i}{\sum_{i=1}^N 1/\sigma_y^i} \end{cases} \quad (23)$$

迭代优化过程中, 在当前解附近一阶泰勒展开, 将非线性优化问题转化为线性优化问题, 根据式(22)得到当前解与最优解之间的偏差 $\Delta\mathbf{X}$, 然后将该偏差与当前解相加作为新的当前解, 进入下一轮迭代求解, 直到问题的解不再变化(趋于稳定)则停止迭代。PnP 位姿估计问题的非线性迭代优化求解需要事先给定一个较为准确的位置和姿态初值, 可采用 PnP 问题的直接求解法得到, 比如 P3P^[2]、EPnP^[3] 等方法。为了保证位姿初值求解的稳定性, 采用随机采样一致性(random sample consensus, RANSAC)方法^[28]剔除关键点检测噪声干扰。

3 实验分析

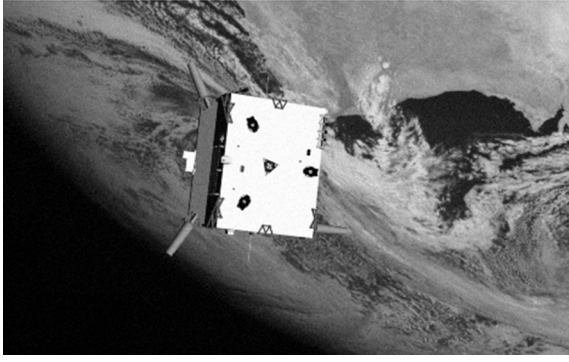
3.1 数据集与评价指标

欧空局和斯坦福大学联合发布的 SPEED 数据集^[26]包括 Tango 卫星的真实图像和仿真图像, 数据集部分图像示例如图 3 所示。仿真图像由计算机渲染仿真生成, 包含深空背景和地球背景两种情形, 真实图像由在实验室搭建模拟太空光照条件的摄像机和卫星半实物仿真平台拍摄得到, 只包含深空背景。数据集由训练集和测试集组成, 训练集包括 5 幅真实图像和 12 000 幅仿真图像, 测试集包括 300 幅真实图像和 2 998 幅仿真图像。



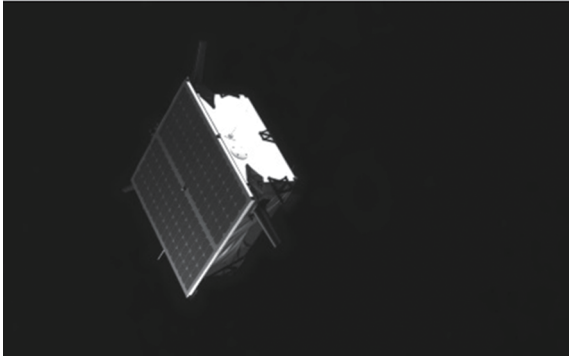
(a) 深空背景卫星仿真图像

(a) Satellite synthetic image with black background



(b) 地球背景卫星仿真图像

(b) Satellite synthetic image with Earth background



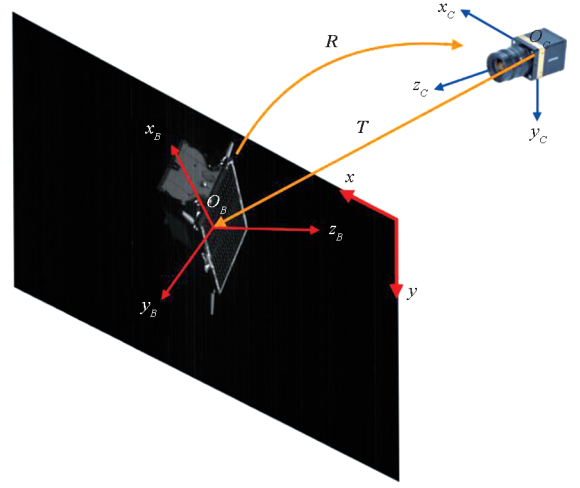
(c) 深空背景卫星真实图像

(c) Satellite real image with black background

图3 SPEED数据集图像示例

Fig. 3 Sample images in SPEED dataset

SPEED 数据集中摄像机坐标系和卫星本体坐标系的定义如图4所示。摄像机坐标系记为 C , 坐标系原点 O_C 为摄像机光心, x_C 轴和 y_C 轴分别与像平面的横轴 x 和纵轴 y 平行, z_C 轴为摄像机光轴并且垂直于 $x_C y_C$ 平面构成直角坐标系。卫星本体坐标系记为 B , 坐标系原点 O_B 位于卫星形心, z_B 轴垂直于太阳能板, x_B 轴和 y_B 轴分别沿太阳能板边缘方向并且与 z_B 轴垂直构成直角坐标系。

图4 SPEED数据集中摄像机坐标系与卫星本体坐标系的定义^[26]Fig. 4 Definition of camera coordinate system and satellite body coordinate system in SPEED dataset^[26]

SPEED 数据集中衡量卫星本体坐标系与摄像机坐标系之间位姿估计精度的评价指标包括姿态和位置两部分。其中, 卫星相对于摄像机的姿态采用四元数表示, 因此该数据集定义的姿态估计误差表示为姿态四元数估计值与真值之间的偏差, 如式(24)所示。

$$E_R = 2 \arccos(|\langle \mathbf{q}, \hat{\mathbf{q}} \rangle|) \quad (24)$$

式中: $\mathbf{q}$ 和 $\hat{\mathbf{q}}$ 分别表示卫星相对于摄像机的姿态估计值和真值的四元数单位向量, $\langle \mathbf{q}, \hat{\mathbf{q}} \rangle$ 表示两个四元数向量的点积。

单目位姿估计求解方法得到罗德里格斯向量表示的姿态关系后, 需转换为四元数再计算式(24)所示的姿态误差。根据四元数运算规则, 单位四元数的逆为该四元数的共轭, 因此, 卫星姿态单位四元数估计值和真值之间的差异可表示为 $\mathbf{e} = \mathbf{q}\hat{\mathbf{q}}^{-1} = \mathbf{q}\hat{\mathbf{q}}^*$ 。 $\mathbf{e}$ 也是一个四元数, 表示卫星从初始姿态先按估计值旋转再按真值旋转回去后的姿态与初始姿态之间的姿态旋转关系, 即为卫星姿态估计误差。四元数 $\mathbf{e}$ 可用轴角法表示, 有 $\mathbf{e} = [\cos(\theta/2), \mathbf{u}\sin(\theta/2)]$, 其中 θ 为旋转角度, $\mathbf{u}$ 为旋转轴向量对应的四元数虚部。只保留四元数 $\mathbf{e}$ 实部, 并取反三角函数即可得到式(24)所示的姿态误差, 其中 E_R 即为姿态估计误差对应的旋转角度 θ 。

位置估计误差表示为卫星本体坐标系原点在摄像机坐标系中的位置向量估计误差与该位置向量真值的模(卫星和摄像机之间距离)之比:

$$E_T = \frac{\|\mathbf{T} - \hat{\mathbf{T}}\|_2}{\|\hat{\mathbf{T}}\|_2} \quad (25)$$

式中: $\mathbf{T}$ 和 $\hat{\mathbf{T}}$ 分别表示三维位置(平移向量)的预测值和真值。

根据 SPEED 数据集中的定义, 将式(24)所示姿态误差和式(25)所示位置误差直接相加作为卫星位姿估计精度评价指标, 如式(26)所示。

$$E = \frac{1}{M} \sum_{i=1}^M (E_R^i + E_T^i) \quad (26)$$

式中: E_R^i 和 E_T^i 分别是第 i 幅测试图像中卫星的姿态和位置误差, M 表示测试图像总数量。评价指标 E 的数值越小表示位姿估计精度越高。

3.2 实验流程与参数设置

卫星单目位姿估计主要流程如图 5 所示: 首先, 对于输入图像, 采用卫星目标检测算法得到卫星外接矩形框; 然后, 根据卫星目标检测结果将卫星图像裁剪并缩放为同一尺寸, 再采用提出的卫星关键点检测与不确定度同步算法, 从裁剪的卫星图像切片中检测卫星关键点, 同步得到卫星关键点坐标及其不确定度; 最后, 在已知摄像机内参、卫星本体坐标系中关键点空间三维坐标以及卫星关键点对应图像二维坐标条件下, 采用加权位姿估计方法求解单目位姿估计的 PnP 问题, 得到卫星本体坐标系相对于摄像机坐标系的位置和姿态参数。

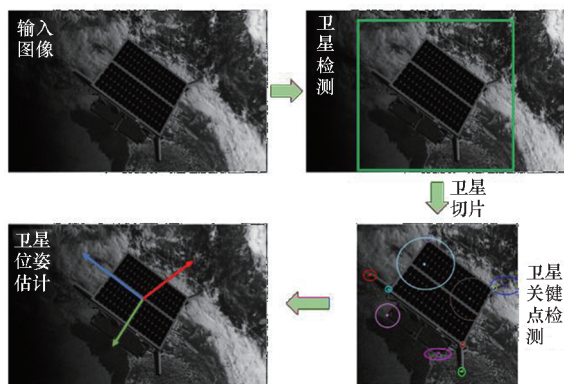


图 5 卫星单目位姿估计流程

Fig. 5 Flowchart of satellite monocular pose estimation

公开发布的 SPEED 数据集中只提供摄像机成像时刻卫星本体坐标系相对于摄像机坐标系的位姿参数, 其中姿态用四元数表示, 位置为三维向量, 并未给出卫星关键点的定义和标注。采用与文献[24]相同的方法, 定义 11 个卫星关键点如图 6 所示。首先, 利用数据集中已知的摄像机成像时刻卫星位姿参数, 选择若干帧图像进行多视图交会测量, 得到卫星关键点在卫星本体坐标系下的三维空间坐标。然后, 再利用数据集中已知的摄像机成像时刻卫星位姿参数, 将已知三维空

间坐标的卫星关键点重投影到图像上, 得到这些卫星关键点的图像坐标, 从而实现了 SPEED 数据集上卫星关键点图像坐标的自动标注, 同时图像上关键点的外接框还可以作为卫星目标检测算法的训练标签。

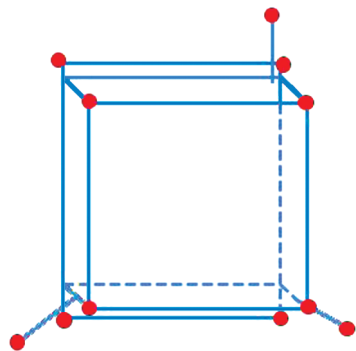


图 6 SPEED 数据集 Tango 卫星 11 个关键点的定义

Fig. 6 Definition of 11 keypoints for Tango satellite in SPEED dataset

卫星目标检测采用与文献[24]相同的 Faster R-CNN^[29], 卫星检测后裁剪的卫星图像切片缩放为 256 像素 × 256 像素输入后续的关键点检测。卫星关键点检测模型的查询向量数量设置为 30, 编/解码器层数设置为 3, 训练时批大小为 12, 迭代轮数(epoch)为 150。

当不考虑卫星关键点图像坐标预测不确定度时, 采用式(14)所示的单目位姿估计方法, 具体流程如下: 先采用 P3P 方法估计位姿初值, 再采用最小二乘迭代求解位姿精确解。当考虑卫星关键点坐标预测不确定度时, 采用式(15)所示的加权单目位姿估计方法, 具体流程如下: 同样先采用 P3P 方法估计位姿初值, 不同的是接下来采用加权最小二乘迭代求解位姿精确解。

3.3 实验结果与分析

实验采用 6 折交叉验证方法评估提出的卫星关键点检测与不确定度同步预测算法在卫星单目位姿估计任务中的性能。将 SPEED 数据集的训练集图像数据随机分为不重复的 6 份, 使用其中 5 份数据进行卫星关键点检测训练, 另外 1 份进行测试, 模型训练和测试重复 6 次, 每次测试数据均不同。然后根据公开数据集 SPEED 中评价准则的定义统计卫星位姿估计误差, 实验结果见表 1。其中, E_R 由式(24)计算得到, 表示根据卫星姿态四元数估计值与真值偏差计算的姿态估计误差; E_T 由式(25)计算得到, 表示摄像机坐标系中卫星位置估计归一化误差, 无量纲; E 由式(26)计算得到, 是 SPEED 数据集中定义的卫星位姿估

表 1 SPEED 数据集 6 折交叉验证实验结果
Tab.1 6-fold cross validation experimental results on the SPEED dataset

方法	评价指标	实验 1	实验 2	实验 3	实验 4	实验 5	实验 6	平均
文献[24]	E_R/rad	0.017 977	0.018 200	0.018 005	0.019 818	0.018 341	0.017 724	0.018 3
	E_T	0.006 710	0.007 965	0.007 310	0.007 638	0.007 910	0.008 451	0.007 7
	E	0.024 687	0.026 164	0.025 316	0.027 455	0.026 251	0.026 175	0.026 0
	E_{t_x}/m	0.004 012	0.004 041	0.003 743	0.004 194	0.004 045	0.004 065	0.004 0
	E_{t_y}/m	0.004 260	0.004 115	0.003 759	0.004 452	0.004 112	0.004 350	0.004 2
	E_{t_z}/m	0.074 536	0.085 007	0.076 809	0.099 166	0.085 854	0.094 081	0.085 9
本文 未加权	E_R/rad	0.012 182	0.012 046	0.013 393	0.011 870	0.011 582	0.012 547	0.012 3
	E_T	0.003 801	0.003 385	0.003 577	0.003 824	0.003 741	0.003 632	0.003 7
	E	0.015 983	0.015 431	0.016 971	0.015 694	0.015 323	0.016 178	0.015 9
	E_{t_x}/m	0.002 486	0.002 518	0.002 560	0.002 668	0.002 473	0.002 651	0.002 6
	E_{t_y}/m	0.002 640	0.002 414	0.002 706	0.002 506	0.002 551	0.002 670	0.002 6
	E_{t_z}/m	0.041 066	0.036 357	0.039 513	0.047 585	0.043 073	0.039 173	0.041 1
本文加权	E_R/rad	0.008 180	0.008 499	0.008 559	0.008 227	0.007 268	0.008 391	0.008 2
	E_T	0.002 833	0.003 087	0.002 863	0.003 126	0.002 870	0.002 782	0.002 9
	E	0.011 013	0.011 585	0.011 422	0.011 353	0.010 138	0.011 173	0.011 1
	E_{t_x}/m	0.001 882	0.001 973	0.001 866	0.002 140	0.001 839	0.001 869	0.001 9
	E_{t_y}/m	0.001 873	0.001 949	0.001 958	0.002 096	0.001 757	0.001 872	0.001 9
	E_{t_z}/m	0.031 622	0.033 222	0.033 442	0.039 266	0.032 814	0.032 082	0.033 7

计算法评价总得分,用来综合衡量位姿估计算法的位置和姿态估计精度; E_{t_x} 、 E_{t_y} 和 E_{t_z} 分别是沿摄像机坐标系三个方向的卫星位置误差,通过估计值和真值直接相减计算绝对值得到。

实验过程中,将文献[24]方法的模型参数设置为与本文方法一致,主要包括卫星目标检测后的图像切片尺寸为 256 像素×256 像素、查询向量数为 30、编/解码器层数为 3 等。采用本文提出的关键点检测与不确定度同步预测方法估计卫星位姿时,首先只使用预测的关键点坐标构建普通最小二乘约束方程直接求解位姿,该方法记为“本文未加权”,然后使用预测的关键点坐标与不确定度构建加权最小二乘约束方程求解位姿,该方法记为“本文加权”。将本文两种求解方法与文献[24]、文献[15]和文献[20]方法的实验结果进行对比,如表 2 所示。其中姿态误差给出了弧度(rad)转换为角度(°)的结果。

当采用同样的模型参数和不加权的最小二乘位姿求解方法时,与文献[24]方法相比,采用本文方法关键点检测结果的卫星位姿估计精度

表 2 SPEED 数据集实验结果对比
Tab.2 Comparison of experimental results on the SPEED dataset

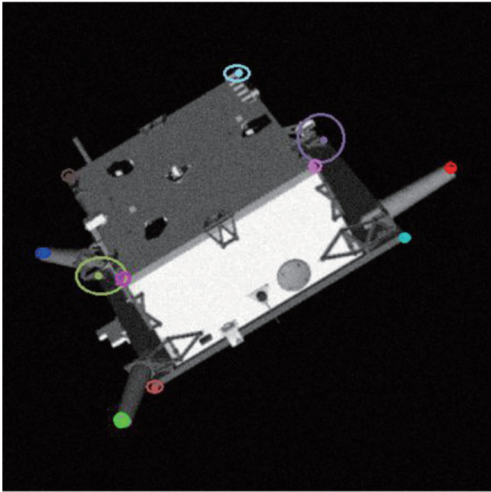
评价 指标	本文 加权	本文 未加权	文献 [24]	文献 [15]	文献 [20]
$E_R/(^{\circ})$	0.469 1	0.703 0	1.051 0	8.430 0	2.872 3
E_R/rad	0.008 2	0.012 3	0.018 3	0.147 1	0.050 1
E_T	0.002 9	0.003 7	0.007 7	—	0.182 3
E	0.011 1	0.015 9	0.026 0	—	0.232 1
E_{t_x}/m	0.001 9	0.002 6	0.004 0	0.055 0	0.010 0
E_{t_y}/m	0.001 9	0.002 6	0.004 2	0.046 0	0.096 0
E_{t_z}/m	0.033 7	0.041 1	0.085 9	0.780 0	0.080 3

显著提升,位姿估计误差评价指标 E 由 0.026 0 (文献[24])下降到 0.015 9(本文未加权)。该结果表明本文提出的关键点检测与不确定度同步预测方法有效提升了卫星关键点检测精度。也就是说,即便位姿求解时不使用关键点坐标预测不确定度,卫星单目位姿估计精度也会有显著提升。

当使用本文方法给出的关键点坐标预测不确

定度和加权最小二乘方法求解位姿时,卫星位姿估计精度进一步显著提升,位姿估计误差评价指标 E 由 0.015 9(本文未加权)下降到 0.011 1(本文加权)。相较于基线方法^[24],位姿估计精度提升一倍以上。实验结果证明了提出的关键点检测与不确定度同步预测方法能够有效提升关键点检测精度以及卫星单目位姿估计精度。

根据图 6 卫星关键点定义,11 个关键点分别是卫星立方体 8 个顶点以及 3 个天线端点。图 7 和图 8 给出了训练过程中(共 150 个 epoch)模型在测试集图像上进行关键点检测与不确定度同步预测的部分实验结果(当 epoch 分别等于 50、100 和 150 时)。图中实心点表示卫星关键点位置真值,实心点外的椭圆表示关键点检测结果,椭圆中心表示关键点预测位置,椭圆尺寸表示关键点二



估计: $[-60.39^\circ \ -65.79^\circ \ -1.68^\circ \ 0.272\ 5\ \text{m} \ 0.021\ 3\ \text{m} \ 6.598\ 0\ \text{m}]$
真值: $[-60.36^\circ \ -65.86^\circ \ -1.51^\circ \ 0.273\ 9\ \text{m} \ 0.022\ 5\ \text{m} \ 6.606\ 0\ \text{m}]$
误差: $[\ 0.02^\circ \ \ 0.07^\circ \ \ 0.17^\circ \ 0.001\ 4\ \text{m} \ 0.001\ 2\ \text{m} \ 0.008\ 0\ \text{m}]$

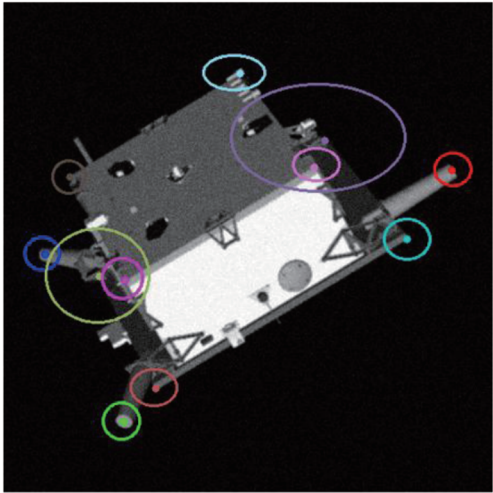
(c) epoch = 150

图 7 卫星关键点检测与不确定度预测结果(示例 1)

Fig. 7 Satellite keypoints detection and uncertainty prediction results (example 1)

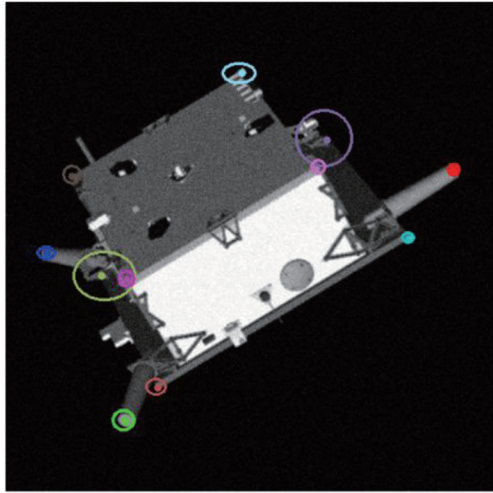
维图像坐标预测不确定度,椭圆横向和纵向主轴越长,表示该方向卫星关键点图像坐标预测的不确定度越大。每幅图给出了卫星相对于摄像机的姿态和位置估计值、真值以及估计误差,为了更加直观,将卫星姿态转化为欧拉角表示。

图 7 和图 8 的实验结果显示,同一幅测试图像中的同一个卫星关键点,从(a)到(c)右椭圆尺寸逐渐减小,表明卫星关键点检测精度随着训练进行在逐渐提高。摄像机成像时,由于遮挡或其他因素,部分卫星关键点不可见或成像不清晰。对于这些关键点,检测时主要由模型依靠图像全局信息综合推测,与清晰可见的关键点相比,其检



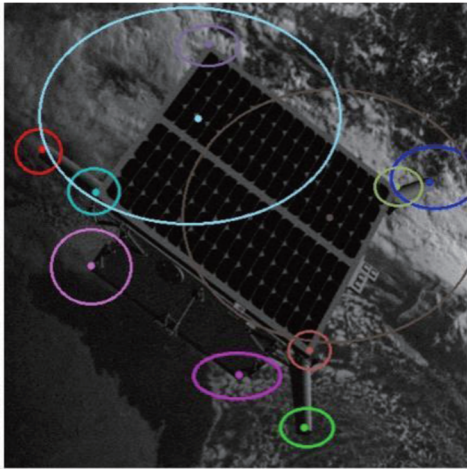
估计: $[-59.68^\circ \ -65.86^\circ \ -0.69^\circ \ 0.269\ 1\ \text{m} \ 0.019\ 5\ \text{m} \ 6.573\ 1\ \text{m}]$
真值: $[-60.36^\circ \ -65.86^\circ \ -1.51^\circ \ 0.273\ 9\ \text{m} \ 0.022\ 5\ \text{m} \ 6.606\ 0\ \text{m}]$
误差: $[\ 0.68^\circ \ \ 0.00^\circ \ \ 0.82^\circ \ 0.004\ 8\ \text{m} \ 0.003\ 0\ \text{m} \ 0.032\ 9\ \text{m}]$

(a) epoch = 50



估计: $[-60.47^\circ \ -65.81^\circ \ -1.52^\circ \ 0.272\ 9\ \text{m} \ 0.021\ 0\ \text{m} \ 6.597\ 4\ \text{m}]$
真值: $[-60.36^\circ \ -65.86^\circ \ -1.51^\circ \ 0.273\ 9\ \text{m} \ 0.022\ 5\ \text{m} \ 6.606\ 0\ \text{m}]$
误差: $[\ 0.11^\circ \ \ 0.06^\circ \ \ 0.01^\circ \ 0.001\ 0\ \text{m} \ 0.001\ 4\ \text{m} \ 0.008\ 6\ \text{m}]$

(b) epoch = 100



估计: $[-123.21^\circ \ -38.74^\circ \ -175.47^\circ \ -0.079\ 8\ \text{m} \ 0.086\ 2\ \text{m} \ 4.036\ 4\ \text{m}]$
真值: $[-122.82^\circ \ -38.81^\circ \ -173.75^\circ \ -0.085\ 4\ \text{m} \ 0.082\ 5\ \text{m} \ 4.055\ 2\ \text{m}]$
误差: $[\ 0.39^\circ \ \ 0.07^\circ \ \ 1.73^\circ \ 0.005\ 6\ \text{m} \ 0.003\ 8\ \text{m} \ 0.018\ 9\ \text{m}]$

(a) epoch = 50

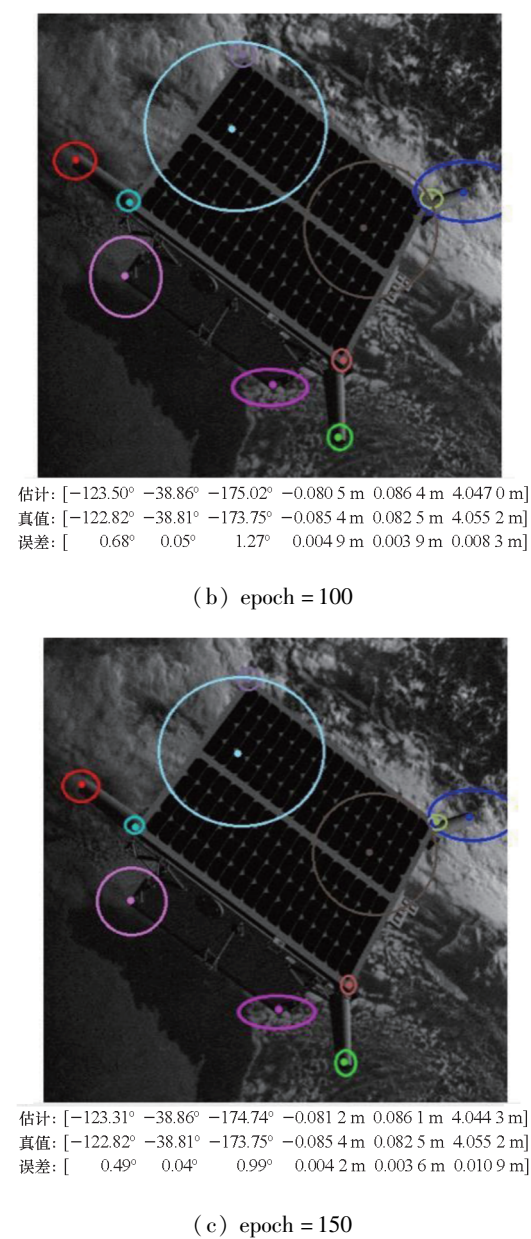


图 8 卫星关键点检测与不确定度预测结果(示例 2)

Fig. 8 Satellite keypoints detection and uncertainty prediction results (example 2)

测精度会相对较差。可以发现,图 7 和图 8 中不可见或成像不清晰的关键点对应的椭圆整体尺寸普遍偏大,意味着本文方法给出的关键点检测不确定度也相对较大,这与理论上预期是一致的,说明了本文方法关键点坐标不确定度预测的准确性和合理性。

对于有些卫星关键点(图 7 右上),训练过程中其对应的不确定度椭圆形状逐渐由扁平趋近于圆。表明对于这些关键点,模型首先关注图像二维坐标其中一个方向的精度,经过不断训练最终实现关键点图像坐标两个方向均达到较好的精度。而对有些卫星关键点(图 8 中下),其对应的不确定度椭圆虽然随着训练在不断减小,但是最

终训练完成后依然是一个较为扁平的椭圆。表明对于这些关键点,其图像坐标的某个方向始终无法达到较好的精度。

4 结论

针对卫星单目视觉位姿估计问题,提出一种关键点检测与不确定度同步预测方法。将关键点检测问题描述为一个高斯分布,高斯分布均值代表关键点坐标,高斯分布标准差代表关键点坐标不确定度,通过构建基于 KL 散度的损失函数,从而实现卫星关键点检测的同时给出关键点二维图像坐标不确定度。在欧空局和斯坦福大学发布的 SPEED 卫星单目位姿估计数据集上开展的充分实验表明,提出的关键点检测与不确定度预测同步方法有效提升了卫星关键点检测精度,而且利用给出的关键点检测不确定度能够自适应调整卫星单目位姿估计的关键点权重,实现单目位姿估计 PnP 问题的加权求解,最终使得卫星单目位姿估计的精度得到显著提升。

参考文献(References)

[1] ANSAR A, DANIILIDIS K. Linear pose estimation from points or lines[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 578–589.

[2] GAO X S, HOU X R, TANG J L, et al. Complete solution classification for the perspective-three-point problem [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(8): 930–943.

[3] LEPETIT V, MORENO-NOGUER F, FUA P. EPnP: an accurate $O(n)$ solution to the PnP problem[J]. International Journal of Computer Vision, 2009, 81: 155–166.

[4] 王刚,施忠臣,尚洋,等. 结合模板匹配和梯度峰值的对角标志自动提取方法[J]. 光学学报, 2018, 38(8): 0815014.

WANG G, SHI Z C, SHANG Y, et al. Automatic extraction of diagonal markers based on template matching and peaks of gradient histogram[J]. Acta Optica Sinica, 2018, 38(8): 0815014. (in Chinese)

[5] DONG S, MA J, SU Z L, et al. Robust circular marker localization under non-uniform illuminations based on homomorphic filtering [J]. Measurement, 2021, 170: 108700.

[6] GARRIDO-JURADO S, MUÑOZ-SALINAS R, MADRID-CUEVAS F J, et al. Automatic generation and detection of highly reliable fiducial markers under occlusion[J]. Pattern Recognition, 2014, 47(6): 2280–2292.

[7] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91–110.

[8] ROSTEN E, DRUMMOND T, ROSTEN E, et al. Machine learning for high-speed corner detection[C]//Proceedings of the 9th European Conference on Computer Vision, 2006: 430–443.

[9] CALONDER M, LEPETIT V, STRECHA C, et al. BRIEF:

- binary robust independent elementary features [C]//Proceedings of the 11th European Conference on Computer Vision, 2010: 778–792.
- [10] RUBLEE E, RABAU D V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF[C]//Proceedings of the International Conference on Computer Vision, 2011: 2564–2571.
- [11] LEUTENEGGER S, CHLI M, SIEGWART R Y. BRISK: binary robust invariant scalable keypoints [C]//Proceedings of the International Conference on Computer Vision, 2011: 2548–2555.
- [12] PAULY L, RHARBAOUI W, SHNEIDER C, et al. A survey on deep learning-based monocular spacecraft pose estimation: current state, limitations and prospects [J]. *Acta Astronautica*, 2023, 212: 339–360.
- [13] PARK T H, MÄRTENS M, JAWAID M, et al. Satellite pose estimation competition 2021: results and analyses [J]. *Acta Astronautica*, 2023, 204: 640–665.
- [14] PROENCA P F, GAO Y. Deep learning for spacecraft pose estimation from photorealistic rendering [C]//Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 2020: 6007–6013.
- [15] SHARMA S, D'AMICO S. Neural network-based pose estimation for noncooperative spacecraft rendezvous [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2020, 56(6): 4638–4658.
- [16] CHEN B, CAO J W, PARRA A, et al. Satellite pose estimation with deep landmark regression and nonlinear pose refinement [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), 2019: 2816–2824.
- [17] WANG Z, CHEN M L, GUO Y L, et al. Bridging the domain gap in satellite pose estimation: a self-training approach based on geometrical constraints [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2023, 60(3): 2500–2514.
- [18] 王梓, 王靖皓, 李杨, 等. 基于轻量级神经网络的非合作目标位姿单目测量 [J]. *航空学报*, 2024, 45(22): 330248.
- WANG Z, WANG J H, LI Y, et al. Non-cooperative target pose estimation from monocular images based on lightweight neural network [J]. *Acta Aeronautica et Astronautica Sinica*, 2024, 45(22): 330248. (in Chinese)
- [19] PARK T H, SHARMA S, D'AMICO S. Towards robust learning-based pose estimation of noncooperative spacecraft [C]//Proceedings of the AAS/AIAA Astrodynamics Specialist Conference, 2019.
- [20] HUAN W X, LIU M M, HU Q L. Pose estimation for non-cooperative spacecraft based on deep learning [C]//Proceedings of the 39th Chinese Control Conference (CCC), 2020: 3339–3343.
- [21] LOTTI A, MODENINI D, TORTORA P, et al. Deep learning for real-time satellite pose estimation on tensor processing units [J]. *Journal of Spacecraft and Rockets*, 2023, 60(3): 1034–1038.
- [22] HUANG H R, SONG B, ZHAO G P, et al. End-to-end monocular pose estimation for uncooperative spacecraft based on direct regression network [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2023, 59(5): 5378–5389.
- [23] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C]//Proceedings of the 16th European Conference, 2020: 213–229.
- [24] 王梓, 孙晓亮, 李璋, 等. 基于 Transformer 模型的卫星单目位姿估计方法 [J]. *航空学报*, 2022, 43(5): 325298.
- WANG Z, SUN X L, LI Z, et al. Transformer based monocular satellite pose estimation [J]. *Acta Aeronautica et Astronautica Sinica*, 2022, 43(5): 325298. (in Chinese)
- [25] WANG Z, ZHANG Z, SUN X L, et al. Revisiting monocular satellite pose estimation with transformer [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2022, 58(5): 4279–4294.
- [26] KISANTAL M, SHARMA S, PARK T H, et al. Satellite pose estimation challenge: dataset, competition design, and results [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2020, 56(5): 4083–4098.
- [27] HE Y H, ZHU C C, WANG J R, et al. Bounding box regression with uncertainty for accurate object detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 2883–2892.
- [28] FISCHLER M A, BOLLES R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography [J]. *Readings in Computer Vision*, 1987: 726–740.
- [29] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137–1149.