中国科学: 技术科学
SCIENTIA SINICA Technologica
techcn.scichina.com
《中国科学》杂志社
SCIENCE CHINA PRESS
论 文

基于无人机视觉的多源位姿融合滑坡监测技术及其应用

张海发, 王一晶, 王浩宇*, 李峥, 左志强

天津大学电气自动化与信息工程学院天津市智能无人集群技术与系统重点实验室, 天津 300072
* E-mail: why2014@tju.edu.cn

收稿日期: 2025-10-11; 接受日期: 2026-02-28; 网络版发表日期: 2026-10-22
国家自然科学基金(批准号: 62173243, 62403348)、天津市自然科学基金(编号: 23JCQNJC01780)、中国博士后科学基金(编号: GZC20241208, 2024M762357)、系统控制与信息处理重点实验室项目(编号: Scip20240116)和天津大学学科交叉中心新兴方向培育专项资助项目

摘要 针对滑坡灾害监测中高精度测量指标与易操作、低成本需求之间的矛盾问题, 本文提出了一种基于多旋翼无人机纯视觉方案的滑坡监测方法. 该方法以"多源位姿融合+全局束调优化"为技术核心, 配合地面可移动靶标以实现低成本、高精度的滑坡监测. 首先, 结合YOLOv8靶标识别与亚像素级特征点定位技术, 精确获取靶标的二维影像坐标; 随后, 通过自适应加权策略, 将无人机位姿与高频视觉位姿进行深度融合; 最终, 在光束法平差框架下对所有相机位姿与靶标三维坐标进行联合优化, 以提高纯视觉系统高程变化测量中的精度. 实验结果表明, 在多角度环绕飞行模式下, 本方法对靶标的绝对位移测量误差能够稳定在3.4~3.8 cm, 显著优于仅依赖无人机位姿或纯视觉运动恢复结构的基线方法. 该方法在保证低成本与高灵活性的前提下, 成功实现了厘米级的高精度形变监测, 为滑坡灾害的早期识别与预警提供了一种经济、高效、可靠的技术途径.

关键词 无人机视觉系统, 滑坡监测, 位姿融合, 光束法平差

1 引言

滑坡作为一种破坏性极强的地质灾害, 对人民生命财产与基础设施安全构成严重威胁, 因此对其进行有效监测与早期预警至关重要. 当前, 无人机遥感技术已成为滑坡监测的重要手段, 但主流的高精度测绘方案多依赖于固定翼无人机搭载大型激光雷达或合成孔径雷达系统[1~3]. 该方案虽能直接获取高精度三维点云, 但存在设备成本高昂、操作复杂、起降场地要求高等问题, 导致其难以大规模推广, 尤其在应急响应和普适化巡检中受到限制[3~5]. 与此同时, 采用小型多旋翼无人机搭载光学相机进行监测, 具有经济实惠、部署灵活、操作简便的显著优势, 是实现滑坡监测低空经济化的理想路径.

滑坡监测的关键在于精确量化山体的高程变化, 这也是纯视觉技术方案面临的核心挑战. 相比于能主动测距的激光雷达, 基于运动恢复结构(structure from motion, SfM)的视觉三维重建在深度和尺度估算上存在固有模糊性[6], 对观测几何构型和算法的鲁棒性要求极高. 此外, 无人机在无风或微风环境下的飞行悬

引用格式: 张海发, 王一晶, 王浩宇, 等. 基于无人机视觉的多源位姿融合滑坡监测技术及其应用. 中国科学: 技术科学
Zhang H F, Wang Y J, Wang H Y, et al. A vision-based UAV landslide monitoring technique using multi-source pose fusion and its application (in Chinese). Sci Sin Tech, doi: 10.1360/SST-2025-0298

© 2026《中国科学》杂志社

停精度也存在上限. 在大多数型号中, 最优的垂直定位精度, 即便在实时动态(real-time kinematic, RTK)差分定位正常工作时也仅为厘米级, 而在标准全球导航卫星系统(global navigation satellite system, GNSS)模式下则为分米级. 这一精度对于大范围测绘已足够, 但对于仅为厘米级的滑坡早期微小形变而言, 无人机自身的位姿数据显然无法直接满足监测要求.

因此, 尽管消费级多旋翼无人机在成本与部署灵活性方面具有明显优势, 但将其直接应用于对测量精度要求极高的滑坡监测场景时, 仍然面临着若干关键性的技术瓶颈, 主要体现在以下三个方面. 首先, 是微小形变信号与平台定位噪声之间的信噪比矛盾. 滑坡孕育与早期变形阶段通常表现为毫米级至厘米级的蠕滑位移, 而消费级无人机即使搭载RTK定位模块, 其在悬停和飞行过程中产生的姿态抖动与定位噪声幅值往往与该形变量级相当. 在此条件下, 若直接基于无人机位姿或影像配准结果进行形变反演, 本质上是在高噪声背景中提取微弱信号, 极易导致误判与虚假报警. 其次, 是垂直高程恢复过程中的几何病态问题. 滑坡监测的核心关注对象通常为坡体沿重力方向的下滑与沉降行为, 然而传统无人机航测多采用近似正射的俯视成像方式, 此时相机光轴方向与重力方向近乎平行. 该几何构型对垂直方向位移的观测敏感度最低, 在纯视觉三维重建或结构光束法平差过程中, 微小的图像匹配误差即可被显著放大为高程方向的系统性漂移, 从而造成平面位置稳定而高程精度不足的问题. 最后, 是纯视觉系统在长时序监测任务中不可避免的尺度漂移与误差累积. 在缺乏高精度控制点或稳定基准等外部强约束情况下, 里程计与视觉系统往往会产生不可逆的尺度不确定性与时间相关误差, 使得重建结果难以直接映射到具有明确地理意义的绝对物理位移量, 限制了其在定量滑坡监测中的应用价值. 综上所述, 在低成本硬件条件难以显著提升的现实背景下, 如何通过算法层面的多源信息融合与几何约束优化, 以突破上述由平台精度与观测几何所带来的限制, 构成了本文研究的核心动机与主要切入点[7,8].

为此, 本文提出了一种面向多旋翼无人机纯视觉系统的滑坡高精度监测方法, 其数据采集方案如图1所示, 即利用机载视觉系统采集人工标靶的图像信息. 该方法以多源位姿融合与全局束调优化为核心, 旨在通过算法实现低成本、高精度的滑坡监测, 并在靶标识别、位姿估计与三维重建等关键环节实现自动化处理. 首先, 基于深度神经网络的目标检测模型自动识别影像中的靶标区域, 随后, 通过亚像素级角点定位算法精确提取靶标特征点在4K影像中的位置; 其次, 系统采用自适应加权策略, 将无人机板载RTK/GNSS提供的全局约束与基于透视n点(perspective-n-point, PnP)算法解算的局部高精度视觉位姿进行融合, 为后续优化提供高质量初值; 最后, 在光束法平差(bundle adjustment, BA)框架下联合优化相机位姿与靶标点三维坐标, 以最小化重投影误差, 从而实现厘米级精度的地表高程与水平位移测量.

2 核心算法原理

2.1 相机成像模型

在摄影测量与计算机视觉领域, 对相机成像过程进行精确的数学建模是实现三维重建与位姿估计的理

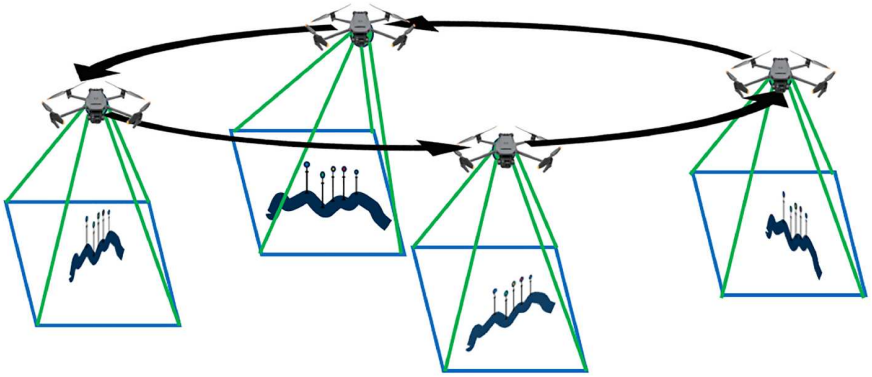


图 1 (网络版彩图)无人机滑坡监测数据采集方案
Figure 1 (Color online) Data acquisition scheme for landslide monitoring.

论基石. 本文采用针孔相机模型来描述物理世界中三维点到二维图像平面上像素点的几何映射关系. 这一映射过程涉及世界坐标系、相机坐标系、图像物理坐标系与像素坐标系之间的一系列几何变换.

首先, 场景中任意三维点在世界坐标系下可表示为$P_w = (X_w, Y_w, Z_w)^{\mathrm{T}}$, 世界坐标系需通过相机的外参转换至相机坐标系$P_c = (X_c, Y_c, Z_c)^{\mathrm{T}}$. 该转换是一个刚体变换, 由旋转矩阵$R$ (3×3)和平移向量t (3×1)共同定义, 描述了相机在世界坐标系中的方位与位置. 随后, 相机坐标系下的三维点依据透视投影原理, 穿过相机光学中心投影至成像平面. 这一过程由相机的内参决定, 这些参数包括以像素为单位的焦距(f_x, f_y)以及成像平面主点坐标(c_x, c_y).

将上述变换过程整合, 可以得到一个从三维世界坐标到二维像素坐标的完整投影方程:

$$s\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} [R|t] \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}, \tag{1}$$

在此方程中, $(u, v, 1)^{\mathrm{T}}$代表三维点最终在图像上的像素坐标. 尺度因子s是一个关键变量, 其物理意义为该三维点在相机坐标系下的深度值Z_c, 反映了投影过程中深度信息的丢失. 左侧的内参矩阵K封装了相机的内部光学特性, 而右侧的外参矩阵$[R|t]$则定义了相机相对于世界坐标系的外部位姿[9].

然而, 理想的针孔模型未能完全刻画真实相机的成像特性, 实际光学镜头存在的几何畸变会使图像产生失真. 畸变通常被分解为由透镜自身形状引起的径向畸变和由相机自身导致的切向畸变. 假设(x, y)为理想无畸变的图像物理坐标, 其到主点的距离为

$$r = \sqrt{x^2 + y^2}. \tag{2}$$

径向畸变可以通过多项式建模为

$$x_{\mathrm{radial}} = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6)x, \tag{3}$$

$$y_{\mathrm{radial}} = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6)y, \tag{4}$$

其中, k_1, k_2, k_3为径向畸变系数. 同时, 切向畸变的数学模型为

$$x_{\mathrm{distorted}} = x + 2p_1 xy + p_2(r^2 + 2x^2), \tag{5}$$

$$y_{\mathrm{distorted}} = y + p_1(r^2 + 2y^2) + 2p_2 xy, \tag{6}$$

式中, p_1, p_2为切向畸变系数. 在实际应用中, 相机的内参矩阵K与畸变系数$(k_1, k_2, k_3, p_1, p_2)$均通过预先的相机标定过程精确测定. 通过对采集到的原始图像应用畸变校正, 可以使其成像特性趋近于理想的针孔模型, 从而为后续高精度三维信息的解算提供可靠的输入数据[10].

2.2 N点透视定位算法

本文采用PnP[11]方法求解相机位姿, 其核心任务是在已知相机内参矩阵K的前提下, 根据一组三维空间参考点在世界坐标系下的坐标$P_{w,i}$及其在图像上的对应二维投影像素坐标p_i, 来求解相机在当前时刻相对于该世界坐标系的位姿, 即相机外参中的旋转矩阵R和平移向量t, 相应的公式为

$$s_i\begin{pmatrix} u_i \\ v_i \\ 1 \end{pmatrix} = K(RP_{w,i} + t), \tag{7}$$

其中, $P_{w,i} = (X_{w,i}, Y_{w,i}, Z_{w,i})^{\mathrm{T}}$是第$i$个世界坐标系下的三维参考点, $p_i = (u_i, v_i)^{\mathrm{T}}$是其对应的像素坐标, s_i是该点在相机坐标系下的深度, 为未知尺度因子. 待求解的未知量为3×3的旋转矩阵R和3×1的平移向量t.

为确保在不同拍摄条件下均能获得稳定可靠的解算结果, 采用结合多种先进PnP算法的策略. 利用高效PnP方法[12,13]将所有三维参考点表示为四个非共面的虚拟控制点$C_{w,j}$的加权线性组合:

$$P_{w,i} = \sum_{j=1}^{4} \alpha_{ij} C_{w,j}, \tag{8}$$

$$\sum_{j=1}^{4} \alpha_{ij} = 1. \tag{9}$$

由于刚体变换保持线性组合关系, 这些点在相机坐标系下同样满足

$$P_{c,i} = \sum_{j=1}^{4} \alpha_{ij} C_{c,j}, \tag{10}$$

其中权重系数α_{ij}可预先计算. 由此, 该问题便从求解n个三维点在相机坐标系下的坐标$P_{c,i}$简化为仅求解4个虚拟控制点的坐标$C_{c,j}$.

2.3 三维坐标解算与全局优化

为给后续的全局优化提供高质量的初值, 需要根据靶标在多张影像上的二维投影点, 并结合已估计的

相机位姿, 来恢复其在三维空间中的坐标. 针对这一任务, 本文采用直接线性变换(direct linear transformation, DLT)[14]进行三角化测量.

该方法的核心在于利用相机投影方程$s_i\boldsymbol{p}_i = P_iP_w$, 其中$P_i$是第$i$个视角的投影矩阵, $\boldsymbol{p}_i$和P_w分别是二维和三维点的齐次坐标. 通过代数变换消除未知的尺度因子s_i, 每个视角的观测都可以转化为关于未知三维点坐标P_w的两个线性方程. 当存在两个或以上视角时, 构成一个形如$AP_w = 0$的齐次线性方程组, 通过奇异值分解求得三维点P_w的坐标. 在本文中, DLT算法用于获取靶标点的初始三维坐标估计.

利用BA方法联合优化相机参数与三维空间点坐标, 尽可能使所有观测达到整体最优一致[15]. 其核心思想是最小化重投影误差. 设第i个相机的内外参数为$(K_i,\ R_i,\ t_i)$, 第j个三维点为X_j, 其在图像上的观测为u_{ij}, 投影函数为$\pi(\cdot)$, 则单个观测的重投影误差可表示为

$$e_{ij} = u_{ij} - \pi\left(K_i, R_i, t_i, X_j\right). \tag{11}$$

在所有相机与所有三维点的观测点集合上, BA的优化目标函数为

$$\min_x = F(x) = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} w_{ij}\left\|e_{ij}\right\|^2, \tag{12}$$

其中x表示所有待优化的参数集合, 包括相机位姿与三维点坐标, w_{ij}为观测权重. 该优化问题具有高度非线性且规模庞大, 直接求解往往不可行, 本文采用Levenberg-Marquardt算法对目标函数在当前估计点处进行线性化近似, 从而得到更新方程:

$$(J^{\mathrm{T}}J + \lambda I)\Delta x = -J^{\mathrm{T}}e, \tag{13}$$

这里J为误差项对参数的雅可比矩阵, $J^{\mathrm{T}}J$近似于目标函数的海森矩阵, λ为阻尼因子. 阻尼因子的作用在于平衡高斯-牛顿法与梯度下降法: 当λ较小时, 更新更接近高斯-牛顿法, 能够以较快速度收敛; 当λ较大时, 更新主要趋向于梯度下降法.

需要指出的是, 式(12)所对应的非线性最小二乘问题采用一阶线性化近似进行迭代求解, 其解算精度在一定程度上依赖于初始参数估计的质量. 当初始位姿与三维点坐标处于最优解的吸引域内时, 线性化误差对最终结果的影响较为有限. 本文通过多源位姿融合为BA优化提供了高质量的初始估计, 并结合Levenberg-Marquardt算法的自适应阻尼机制, 有效提高了优化过程的收敛稳定性与鲁棒性.

3 基于多源位姿融合的滑坡监测方法

3.1 系统整体框架

为实现基于纯视觉的滑坡高精度监测以及数据高效处理, 本文设计了一套完整的数据处理与分析框架, 其整体流程如图2所示. 该框架以多源位姿融合与全局优化为核心, 涵盖了从原始数据采集、位姿估计、三维重建到精度评估的全过程. 系统接收无人机采集的多视角图像与GNSS/RTK位姿作为双重数据输入, 通过一系列精细化的处理与优化步骤, 最终输出高精度的地面靶标三维坐标, 为多时相形变分析提供可靠依据.

系统的处理流程首先从数据预处理开始. 对于位姿数据流, 系统执行坐标系转换, 将原始的WGS-84大地坐标投影至本地笛卡尔坐标系, 并依据时间戳与图像数据进行严格同步. 同时, 对位姿数据进行异常值检测, 剔除因信号失锁或多路径效应引起的粗差. 具体而言, 基于无人机最大飞行速度$v_{\max}$设定阈值计算相邻时间戳Δt内的位移量ΔP, 若在Δt内$\|\Delta P\| > \Delta t v_{\max} + 3\sigma$, 其中$\sigma$为定位噪声标准差, 则判定为异常跳变点并予以剔除. 对于图像数据流, 由于实验所用为定焦

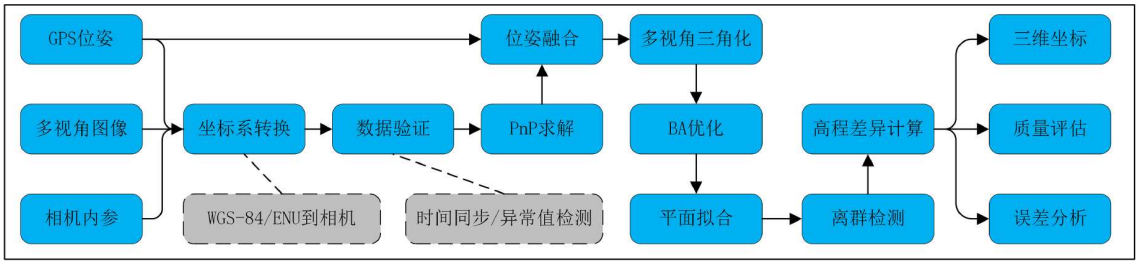


图 2 (网络版彩图)基于多源位姿融合的滑坡监测系统流程图
Figure 2 (Color online) Framework of the landslide monitoring system with multi-source pose fusion.

镜头, 其光学特性在飞行过程中保持稳定, 因此本文对所有视角的影像采用统一的去畸变映射, 利用预先标定的相机内参对图像进行畸变校正, 并采用亚像素级角点检测算法, 提取每张影像中地面靶标的二维像素坐标.

核心环节在于相机位姿的高精度解算与三维重建. 系统并行处理两条位姿信息流: 一方面, 预处理后的无人机位姿数据提供了全局但精度有限的定位基准; 另一方面, 系统利用靶标的三维模型点及其在单张影像上的二维投影, 通过PnP求解模块实时解算出高频、高相对精度的局部视觉位姿. 随后, 将上述两种位姿信息依据信号质量和解算稳定性进行自适应加权融合, 生成兼具全局稳定性和局部精确性的优化位姿.

基于融合后的高精度位姿, 系统进入三维重建与优化阶段. 首先通过DLT三角化对靶标点进行计算, 获得其三维坐标的初始估计. 而后, 将所有融合位姿和初始三维点作为输入, 送入BA优化模块. 该模块通过最小化重投影误差, 对所有相机位姿和靶标点的三维坐标进行联合优化, 系统性地消除累积误差, 输出高精度、全局一致的三维点云.

最后, 系统对优化后的靶标点云进行平面拟合, 并从中计算出靶标的精确中心点坐标, 以进一步提升测量精度. 通过离群点检测剔除无效解算结果后, 系统进行最终的质量评估, 输出各靶标点的三维坐标及相应的精度指标. 通过对比不同监测周期获取的坐标成果, 即可精确量化滑坡体的三维形变. 整个框架通过模块化的设计, 实现了从数据到结果的全自动化处理, 确保了在复杂环境下的监测精度与鲁棒性.

3.2 数据采集与预处理

这一阶段旨在获取几何构型优良的多视角影像及同步的无人机位姿信息, 并将其转换为后续算法所需的标准格式.

为验证算法在不同观测条件下的性能, 数据采集包含正射与多角度环绕两种飞行模式. 在正射模式下, 无人机在目标区域上空以恒定高度飞行, 相机镜头垂直朝下, 按照预设航线进行拍摄. 为保证模型重建的完整性, 航线重叠度设置不低于80%. 该模式主要用于大范围区域的快速测绘与基准模型的建立. 而在多角度环绕模式下, 无人机围绕靶标布设区域进行多角度、不同高度的环绕飞行拍摄, 确保观测视角能够覆盖30°~60°的范围, 且相机间的基线长度与目标距离之比不小于0.2. 该模式通过构建几何约束, 是实现高精度三维解算与微小形变测量的关键.

在获取原始数据后, 需进行一系列预处理操作, 以消除系统误差并统一数据基准. 首先是对无人机记录的位姿数据进行坐标系转换. 原始的GNSS/RTK数据通常以WGS-84大地坐标系下的经纬度形式提供. 为便于后续的三维空间计算并消除地球曲率的影响, 系统首先将其通过通用横轴墨卡托投影转换为平面直角坐标. 并以首次采集影像时无人机的位置为原点, 建立一个局部相对坐标系, 并将所有位姿数据转换至该局部框架下.

3.3 位姿估计与全局优化

为实现低成本、高可靠性的滑坡监测, 本文方法对无人机相机位姿及地面靶标三维坐标进行高精度估计. 该过程采用一个两阶段策略: 首先通过多源信息融合生成一个高质量的初始估计, 然后通过BA进行最终的联合优化.

在第一阶段, 即初始位姿融合中, 核心目标是充分利用无人机板载的GNSS/RTK位姿与基于PnP算法解算的视觉位姿, 通过优势互补, 获得远超单一信息源的精度与鲁棒性. GNSS/RTK数据提供了一个无长期漂移的全局定位基准, 但其更新频率较低且易受信号干扰; 而PnP视觉位姿则具有极高的更新频率, 能够提供平滑且高相对精度的局部几何信息.

对于相机的位置分量, 采用线性加权的方式进行融合. 设由GNSS/RTK模块提供的相机位置向量为P_{GPS}, 通过PnP算法解算出的视觉位置向量为P_{PnP}, 则融合后的位置P_{fused}可表示为

$$P_{\mathrm{fused}} = \alpha P_{\mathrm{GPS}} + (1-\alpha) P_{\mathrm{PnP}}. \tag{14}$$

在具体部署时, 权重α取0.2, 该分配策略充分信任由视觉数据解算出的高频PnP结果, 同时利用无人机位姿数据作为全局约束以抑制漂移.

对于相机的旋转分量, 为确保旋转变化的平滑性与数学上的最优性, 采用SLERP球面线性插值方法进行融合[16]. 设由无人机位姿和PnP算法解算出的相机姿态, 分别用单位四元数q_{GPS}和q_{PnP}表示, 其夹角为$\Omega = \arccos q_{\mathrm{GPS}} q_{\mathrm{PnP}}$. 则融合后的旋转$q_{\mathrm{fused}}$可通过SLERP公式计算得出:

$$q_{\text{fused}} = \frac{\sin(1-\beta)\Omega}{\sin\Omega} q_{\text{GPS}} + \frac{\sin\Omega\beta}{\sin\Omega} q_{\text{PnP}}, \tag{15}$$

其中, PnP的权重β取0.8. 该公式保证了融合旋转在四维球面上沿着最短路径以恒定角速度进行插值.

在第二阶段, 即全局优化中, 系统将第一阶段获得的融合位姿以及通过DLT三角化得到的靶标点三维坐标初始估计, 一同送入BA框架以进行最终的联合精化.

BA的核心是最小化重投影误差. 在本系统中, 该过程的具体实现步骤如下. 首先, 构建一个包含所有待优化参数的状态向量x, 该向量包括了全部m个相机位姿和全部n个靶标点的三维坐标, 可表示为

$$x = [a_1, \ldots, a_m, X_1, \ldots, X_n]^{\mathrm{T}}, \tag{16}$$

其中a_i代表第i个相机位姿的参数, X_j代表第j个靶标点的三维坐标.

随后, 系统依据式(12)构建非线性最小二乘问题, 并采用Levenberg-Marquardt算法进行迭代求解. 在每一次迭代中, 系统都将执行一个包含计算、求解和评估的完整闭环流程. 首先, 对于当前的参数估计值x, 计算代价函数中每一项重投影误差e_{ij}关于相应相机位姿参数a_i和三维点坐标X_j的偏导数, 从而构建出高维但高度稀疏的雅可比矩阵J. 其次, 根据式(13)构建并求解正规方程$(J^{\mathrm{T}}J+\lambda I)\Delta x = -J^{\mathrm{T}}e$, 以计算出最优的参数更新增量$\Delta x$.

最后, 系统使用该增量更新状态向量, 并重新计算总的重投影误差. 本次更新是否被接受, 取决于其是否能减小代价函数的值: 如果误差减小, 则接受本次更新, 并减小阻尼因子λ加速收敛; 反之, 如果误差增大, 则拒绝本次更新, 并增大λ以使更新步骤更加保守.

3.4 靶标设计与自动化识别

为实现图像特征点的高效选取, 本文设计了一套从靶标物理设计到算法实现的完整流程, 旨在确保在复杂野外环境下的识别鲁棒性与定位精度.

在靶标物理设计层面, 选用具有明确几何约束的棋盘格作为地面监测靶标. 棋盘格因其内部角点突出、拓扑结构稳定, 已拥有成熟的计算机视觉检测算法. 为确保在无人机常规作业高度(如50~100 m)下仍能清晰成像, 靶标物理尺寸设计为100 cm×100 cm, 内部由高对比度的黑白相间正方形格子组成, 每个格子的边长为25 cm. 这种大尺寸、高对比度的设计, 为后续的图像处理提供了高质量的原始信号, 能够有效抵抗光照变化和部分遮挡等不利因素的干扰. 具体来说, 所采用的棋盘格靶标单个包含25个可用特征点, 红白航测靶标单个包含6个特征点, 在实际滑坡监测应用中通常会在同一区域布设多个靶标, 从而在特征点数量与空间分布上形成充足的观测冗余.

在自动化识别算法层面, 为兼顾大场景下的检测效率与高精度的定位要求, 本文提出了一种层次化的靶标检测与定位策略. 该策略的第一阶段为基于深度学习的候选区域提取. 考虑到无人机航拍影像具有背景复杂、目标尺度变化大等特点, 本文采用先进的YOLOv8[17]目标检测模型对靶标进行快速的初步定位. 为保证模型的鲁棒性与泛化能力, 构建了一个包含多种光照、视角及背景条件的专用数据集, 并通过旋转、缩放及噪声注入等数据增强技术对其进行扩充. 训练后的模型能够高效地处理4K分辨率的影像, 并以边界框的形式输出所有潜在靶标的候选区域, 从而极大地缩小了后续精细化搜索的计算范围.

该策略的第二阶段为基于几何约束的亚像素特征点定位. 在YOLOv8提供的每个候选区域内, 系统利用靶标自身的几何特性进行精确的角点提取. 首先, 采用自适应阈值二值化方法以增强图像的局部对比度, 凸显棋盘格的结构边缘. 随后, 通过一种基于角点拓扑结构约束的搜索算法, 在二值化图像中识别并排序所有内部角点, 以使最终的定位精度达到像素级别.

4 实验与结果分析

4.1 实验设计

本次实验在一处地形开阔、无高大遮挡物的场地进行, 以确保无人机飞行安全与GNSS信号的良好接收, 部分采集数据如图3所示. 数据采集平台为大疆M4E多旋翼无人机, 如图4所示, 其搭载的4K分辨率光学相机为数据采集提供了丰富的细节信息. 在靶标选择上, 本方法可适配多种靶标, 为便于进行高精度的几何验证, 本次实验以具有明确拓扑结构和高对比度的100 cm×100 cm棋盘格为主要的测量样本.

为建立精确的测量基准并模拟真实的地表位移, 将多个棋盘格靶标分别固定在可三维调节的精密支架上. 在两次数据采集任务之间, 通过调节支架, 可以使



图 3 (网络版彩图)多角度环绕飞行部分实拍图
Figure 3 (Color online) A field photo of the multi-angle orbiting flight.

靶标在垂直和水平方向上产生预设的精确位移. 在每次数据采集前后, 均对各靶标的三维坐标进行独立测量, 这些测量结果及其精确的位移量, 将作为评估本方法解算精度的真值.

为全面评估本方法在不同观测几何构型下的性能, 如表1所示, 数据采集包含了两种飞行模式. 第一种为正射航线飞行, 无人机在靶标区域上空恒定50 m的高度, 以80%的重叠度沿规划航线飞行. 第二种为多角度环绕飞行, 无人机以靶标布设区为中心, 30~60 m的变高飞行以及−45°~−60°的倾斜观测视角, 以构建几何约束, 该模式旨在验证本方法在精细化形变监测中的最高精度. 为了减少数据冗余并提高飞行作业效率, 本文对环绕路径进行了优化设计, 放弃了传统的全频次连续抓拍模式, 采用关键视点规划策略. 具体而言, 依据摄影测量中的最佳交会角原理, 将环绕拍摄的方位角间隔设定为10°~15°, 并确保相邻视角的基高比不小于0.2. 所有飞行任务均在天气晴朗、风力微弱的条件下进行, 以最大程度减小环境因素对数据质量的干扰. 与此同时, 本文采用高快门速度和高对比度棋盘格靶标, 以提高亚像素级特征点定位在实际采集条件下的稳定性.



图 4 (网络版彩图)大疆M4E多旋翼无人机
Figure 4 (Color online) The DJI M4E multi-rotor UAV.

4.2 评价指标

为对实验结果进行客观、定量的评价, 本文引入了以下关键指标, 分别从绝对定位精度、位移测量精度和内部几何一致性三个维度来衡量算法的性能. 其中, 绝对定位精度指标用于评估解算三维坐标相对于外部真值的偏差水平; 位移测量精度指标用于检验多时相条件下的真实形变反演能力; 内部几何一致性指标则从优化拟合角度反映模型与观测的匹配程度. 通

表 1 不同飞行模式下的采集参数配置与路径规划对比

Table 1 Comparison of acquisition parameters and path planning for different flight modes

飞行模式	飞行高度 (m)	云台俯仰角 (°)	航向重叠率/视角间隔
正射航线飞行	50	−90	80%
多角度环绕飞行	30~60	−45~−60	间隔10°~15°

过上述指标的联合分析，可更全面地对不同方法在精度与稳定性方面进行对比评价.

首先，为评估算法解算出的靶标三维坐标与全站仪测量的真值之间的整体偏差，本文采用三维坐标均方根误差作为核心指标. 设算法解算出的第i个靶标点的三维坐标为$P_{\mathrm{est},i}=(X_{\mathrm{est},i},Y_{\mathrm{est},i},Z_{\mathrm{est},i})$，其对应的真值坐标为$P_{\mathrm{true},i}=(X_{\mathrm{true},i},Y_{\mathrm{true},i},Z_{\mathrm{true},i})$，总靶标点数为$N$. 则三维均方根误差定义为

$$\mathrm{RMSE}_{3\mathrm{D}}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left\|P_{\mathrm{est},i}-P_{\mathrm{true},i}\right\|^2}. \tag{17}$$

其次, 考虑到滑坡监测对高程变化的敏感性, 本文进一步将$\mathrm{RMSE}_{3\mathrm{D}}$分解为水平误差$\mathrm{RMSE}_{XY}$和垂高程误差$\mathrm{RMSE}_{Z}$，以更细致地分析误差在不同方向上的分布特性, 它们的具体形式如下:

$$\mathrm{RMSE}_{XY}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}[(X_{\mathrm{est},i}-X_{\mathrm{true},i})^2+(Y_{\mathrm{est},i}-Y_{\mathrm{true},i})^2]}, \tag{18}$$

$$\mathrm{RMSE}_{Z}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(Z_{\mathrm{est},i}-Z_{\mathrm{true},i})^2}. \tag{19}$$

最后，为衡量BA优化结果的内部几何一致性，本文引入了平均重投影误差. 该指标不依赖于外部真值, 它表示所有优化后的三维点，根据优化后的相机位姿重新投影回图像平面时，其预测像素位置与原始观测像素位置之间的平均偏差. 该值越小，表明最终的三维模型与所有原始二维图像的吻合程度越高，优化结果越理想. 其计算公式为

$$E_{\mathrm{reproj}}=\frac{1}{K}\sum_{i=1}^{m}\sum_{j=1}^{n}\omega_{ij}\left\|u_{ij}-\pi(K_i,R_i,t_i,X_j)\right\|, \tag{20}$$

其中K为总的有效观测数量，u_{ij}为三维空间点在第i张影像上的实际观测像素坐标，$\pi(K_i,\ R_i,\ t_i,\ X_j)$则是利用优化后的相机内外参$(K_i,\ R_i,\ t_i)$和三维点坐标$X_j$计算出的理论投影坐标.

4.3 实验结果与分析

由于每次独立的三维重建任务均在各自的虚拟坐标系中完成, 为对两次数据进行比较, 首要步骤是进行坐标系对齐. 本文利用场景中未移动的稳定靶标作为控制点，通过刚体变换将第二期重建的点云对齐至第一期的坐标系下，以建立统一的比较基准. 后续所有误差分析均在此对齐后的坐标系中进行.

系统在两次独立测量中重建场景几何形态的一致性, 是能够精确测量微小位移的基础. 本文通过计算对齐后稳定靶标的三维坐标RMSE来衡量系统的重复性与一致性. 每种飞行模式下均进行了21次独立飞行与重建. 结果如表2所示，本文将所提方法与两种基准方法进行了对比. 由表可知，无论采用何种数据处理方法，多角度环绕飞行的重建精度均系统性地优于传统的正射航线飞行. 在同样的实验条件下，多角度环绕飞行可将三维重建误差从9.45 cm减少至3.20 cm. 其根本原因在于, 正射航线的观测视角单一, 导致在深度方向的三角测量基线较短，解算三维点位的几何约束较弱. 相反，多角度、多高度的环绕飞行能够从不同视角观测同一目标，构建出更长的三角测量基线，从而极大地增强了光束法平差解算过程中的几何约束强度, 有效降低了三维点位的深度不确定性.

为了说明本文方法的有效性，我们将其与目前主流的仅依赖GNSS/RTK位姿的直接地理配准方法[18]和SfM方法[9]进行了对比. 注意到不同的数据融合策略对最终精度有显著影响，SfM方法因缺乏绝对尺度基准和全局位置约束，其解算结果存在固有的尺度模糊性和累积漂移问题，导致其坐标误差达到分米级，无法满足变形监测的要求. 引入GNSS/RTK位姿作为全局约束后，可有效消除尺度漂移，将精度提升至厘米级. 然而, 其精度上限受限于机载GNSS/RTK模块自身的噪声水平和数据更新频率. 本文提出的多源融合策略，在利用无人机位姿提供绝对位置锚点的同时，引入了高频的PnP来估计相机帧间的相对运动. 这种高

频的相对位姿信息为BA优化提供了更为平滑和精确的位姿初值, 使得非线性优化更容易收敛到更优的解, 最终体现为更低的重投影误差和更高的三维坐标一致性.

在评估完系统的静态稳定性后, 仍需检验其对真实地表位移的监测能力. 通过将各方法解算出的靶标位移与全站仪测量的高精度真值进行比较, 计算绝对位移误差.

对于动态测量, 在每一种预设位移条件下均进行了共计17次独立采集与解算, 结果如表3所示, 在采用多角度飞行策略时, 本文方法的位移测量误差能够维持在3.4~3.8 cm的区间内, 且该误差水平并未随着位移量从5.00 cm增加到12.50 cm而发生显著变化.

上述结果表明, 本系统的测量误差主要来源于其固有的三维点位重建不确定性, 而非位移量级本身. 换言之, 测量误差是两次独立重建任务中目标点定位误差的综合体现. 只要场景的几何特征与观测条件保持稳定, 系统对目标点的定位精度就不会改变, 因此其位移测量误差也保持稳定. 这对于滑坡监测至关重要, 因为它意味着该方法对从早期微小变形到加速阶段的较大变形均能提供一个稳定、可预期的误差范围.

相比之下, 其他基线方法的表现则与其静态精度直接相关. 仅依赖GNSS/RTK的方法, 其位移误差略高, 这与其相对较高的点位不确定性相符. 而SfM方法, 其静态重建误差远大于真实的位移信号, 导致其位移测量结果完全失效, 无法从中提取任何有效信息. 此外, 从两种不同类型靶标的测量结果分析, 误差差异微小, 表明本文中采用的靶标自动识别与中心定位算法具有足够的鲁棒性, 未引入显著的系统偏差.

表 2 不同方法在固定靶标上的重复性精度

Table 2 Repeatability accuracy of different methods on stationary targets

飞行模式	方法	$RMSE_{XY}$ (cm)	$RMSE_{Z}$ (cm)	$RMSE_{3D}$ (cm)	E_{reproj} (pixel)
正射航线飞行	本文方法(多源融合+BA)	4.83	8.12	9.45	1.92
	GNSS/RTK位姿+BA[18]	6.24	8.91	10.88	3.46
	纯视觉SfM[9]	25.41	38.79	46.37	2.42
多角度环绕飞行	本文方法(多源融合+BA)	1.88	2.59	3.20	1.12
	GNSS/RTK位姿+BA[18]	2.52	3.62	4.41	2.86
	纯视觉SfM[9]	16.82	24.26	29.52	1.82

表 3 不同方法在移动靶标上的绝对位移误差对比

Table 3 Comparison of absolute displacement errors on moving targets for different methods

飞行模式	靶标类型	真实位移 (cm)	本文方法误差 (cm)	仅GNSS/RTK误差[18] (cm)	纯视觉SfM误差[9] (cm)
正射航线飞行	棋盘格靶标	5.00	9.98	11.63	48.77
		8.25	10.04	11.55	48.62
		12.50	9.89	11.68	48.78
	红白航测靶标	5.00	10.09	11.72	49.92
		8.25	10.05	11.99	49.99
		12.50	10.11	12.01	50.08
多角度环绕飞行	棋盘格靶标	5.00	3.45	4.92	31.87
		8.25	3.42	4.98	31.09
		12.50	3.66	4.82	31.55
	红白航测靶标	5.00	3.77	5.11	32.67
		8.25	3.69	5.12	32.59
		12.50	3.73	5.33	32.96

5 结论与讨论

针对传统滑坡监测方案往往依赖合成孔径雷达或激光雷达等高成本传感设备、系统复杂度较高，而低成本消费级无人机又面临定位精度不足的挑战，本文提出一种以多源位姿融合与全局束调优化为核心的高精度滑坡监测方法. 该方法通过算法层面的设计, 充分挖掘低成本纯视觉硬件的潜力，在不依赖合成孔径雷达或激光雷达等高成本传感器的前提下，实现厘米级地表形变监测.

通过自适应加权融合无人机板载位姿信息与PnP算法解算的高频视觉位姿，该方法为全局束调整提供了高质量的初始值，从而实现了高精度的三维重建. 本研究同样揭示了观测几何构型是决定监测精度的关键因素, 相较于传统的正射航线飞行, 多角度、多高度的环绕飞行模式能够构建更强的几何约束，将重建与位移测量精度提升60%以上. 在此优化条件下, 静态场景的重复性三维重建精度可达3.20 cm，并且对5.00~12.50 cm范围内的真实地表位移，绝对测量误差能够稳定在3.4~3.8 cm的区间内. 此外, 通过结合深度学习目标检测与亚像素角点定位算法，本文方法在靶标识别、位姿估计与三维重建等关键环节实现了高度集成的数据处理流程, 有效降低了人工干预需求, 提升了监测效率与结果稳定性.

本文研究表明，通过观测几何设计与多源位姿融合算法的协同优化，可以在保持系统低成本与高灵活性的同时，实现满足实际工程需求的高精度滑坡监测, 为普适化、经济高效的滑坡灾害监测与早期预警提供一种可行的技术路径.

参考文献

1 Yin Y P, Liu X J, Zhao C Y, et al. Multi-dimensional and long-term time series monitoring and early warning of landslide hazard with improved cross-platform SAR offset tracking method. Sci China Tech Sci, 2022, 65: 1891–1912

2 Casagli N, Intrieri E, Tofani V, et al. Landslide detection, monitoring and prediction with remote-sensing techniques. Nat Rev Earth Environ, 2023, 4: 51–64

3 Zhou X P, Ye T, Huang X C. Determinations of the early warning threshold for creep landslides and their applications (in Chinese). Sci Sin Tech, 2022, 52: 1802–1818 [周小平, 叶腾, 黄小城. 蠕变型滑坡阈值确定方法及工程应用. 中国科学: 技术科学, 2022, 52: 1802–1818]

4 Zhang K X. Review on geological disaster monitoring and early warning system based on “3S” technology in China (in Chinese). Chin J Geol Hazard Control, 2020, 31: 1–11 [张凯翔. 基于“3S”技术的地质灾害监测预警系统在我国应用现状. 中国地质灾害与防治学报, 2020, 31: 1–11]

5 Xu Q, Guo P Y. Mechanisms and characteristics of landslide deformation (in Chinese). Sci Sin Tech, 2025, 55: 1222–1234 [许强, 郭朋瑜. 滑坡变形机理与特征研究. 中国科学: 技术科学, 2025, 55: 1222–1234]

6 Zhang Q, Huang G W, Yang C S. Precision space observation technique for geological hazard monitoring and early warning (in Chinese). Acta Geod Cartogr Sin, 2017, 46: 1300–1307 [张勤, 黄观文, 杨成生. 地质灾害监测预警中的精密空间对地观测技术. 测绘学报, 2017, 46: 1300–1307]

7 Carlà T, Tofani V, Lombardi L, et al. Combination of GNSS, satellite InSAR, and GBInSAR remote sensing monitoring to improve the understanding of a large landslide in high alpine environment. Geomorphology, 2019, 335: 62–75

8 Wang D, Huang G, Du Y, et al. Stability analysis of reference station and compensation for monitoring stations in GNSS landslide monitoring. Satell Navig, 2023, 4: 29

9 Szeliski R. Computer Vision: Algorithms and Applications. Cham: Springer Nature, 2022.

10 Zhang Z. A flexible new technique for camera calibration. IEEE Trans Pattern Anal Machine Intell, 2000, 22: 1330–1334

11 Gao X S, Hou X R, Tang J L, et al. Complete solution classification for the perspective-three-point problem. IEEE Trans Pattern Anal Machine Intell, 2003, 25: 930–943

12 Lepetit V, Moreno-Noguer F, Fua P. EPnP: An accurate O(n) solution to the PnP problem. Int J Comput Vis, 2009, 81: 155–166

13 Terzakis G, Lourakis M. A consistently fast and globally optimal solution to the perspective-n-point problem. In: Proceedings of the European Conference on Computer Vision. 2020. 478–494

14 Xue Y, Zhang S, Zhou M, et al. Novel SfM-DLT method for metro tunnel 3D reconstruction and Visualization. Underground Space, 2021, 6:

134–141
15 Liu Z, Liu X, Zhang F. Efficient and consistent bundle adjustment on LiDAR point clouds. IEEE Trans Robot, 2023, 39: 4366–4386
16 Wang Y, Gui S, Tian Z, et al. A high-precision GNSS SAR imaging fusion method utilizing optimally matched satellites calculated by CRLB. IEEE Trans Geosci Remote Sens, 2025, 63: 1–14
17 Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016. 779–788
18 Štroner M, Urban R, Seidl J, et al. Photogrammetry using UAV-mounted GNSS RTK: Georeferencing strategies without GCPs. Remote Sens, 2021, 13: 1336

A vision-based UAV landslide monitoring technique using multi-source pose fusion and its application

ZHANG HaiFa, WANG YiJing, WANG HaoYu, LI Zheng & ZUO ZhiQiang

Tianjin Key Laboratory of Intelligent Unmanned Swarm Technology and Systems, School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

To address the contradiction between high-precision measurement requirements and the demand for ease of operation and low cost in landslide monitoring, this paper proposes a landslide monitoring method based on a purely visual approach using a multi-rotor unmanned aerial vehicle (UAV). The method takes “multi-source pose fusion + global bundle adjustment optimization” as its technical core, combined with ground mobile targets, to achieve low-cost and high-precision landslide monitoring. First, YOLOv8 target detection and sub-pixel feature point localization techniques are employed to accurately obtain the two-dimensional image coordinates of the targets. Then, through an adaptive weighting strategy, UAV poses are deeply fused with high-frequency visual poses. Finally, within a bundle adjustment framework, all camera poses and three-dimensional target coordinates are jointly optimized to improve the accuracy of elevation change measurements in purely visual systems. Experimental results demonstrate that, under a multi-angle circular flight mode, the absolute displacement measurement error of targets can be stably maintained within 3.4–3.8 cm, which is significantly superior to baseline methods relying solely on UAV poses or pure structure-from-motion approaches. This method, while ensuring low cost and high flexibility, successfully achieves centimeter-level high-precision deformation monitoring, providing an economical, efficient, and reliable new technological approach for the early identification and warning of landslide disasters.

UAV vision system, landslide monitoring, pose fusion, bundle adjustment

doi: 10.1360/SST-2025-0298