



测绘科学  
*Science of Surveying and Mapping*  
ISSN 1009-2307, CN 11-4415/P

## 《测绘科学》网络首发论文

题目: 无人机视频影像的准实时融合算法  
作者: 耿中元, 袁继英, 李英成, 孙新博, 朱祥娥, 冯亮  
收稿日期: 2018-07-16  
网络首发日期: 2019-06-24  
引用格式: 耿中元, 袁继英, 李英成, 孙新博, 朱祥娥, 冯亮. 无人机视频影像的准实时融合算法[J/OL]. 测绘科学.  
<http://kns.cnki.net/kcms/detail/11.4415.P.20190621.1331.016.html>



**网络首发:** 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

**出版确认:** 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

# 无人机视频影像的准实时融合算法

耿中元<sup>1,2</sup>, 袁继英<sup>3</sup>, 李英成<sup>1,2,4</sup>, 孙新博<sup>1,2</sup>, 朱祥娥<sup>1,2</sup>, 冯 亮<sup>1,2</sup>

(1.航空遥感技术国家测绘地理信息局重点实验室, 北京 100039;

2.中测新图(北京)遥感技术有限责任公司, 北京 100039;

3.内蒙古自治区土地调查规划院, 呼和浩特 010010;

4.中国测绘科学研究院, 北京 100830)

**摘要** 无人机视频影像可以实时地融合到地理信息系统中, 但是无人机的位置和姿态数据存在较大误差, 导致相邻的视频影像的同名地物点并不重合。针对该问题, 该文提出准实时的无人机视频影像融合算法, 采用基于滑动窗口的光束平差方法来优化相机的外方位元素, 这种方法比全部视频影像做整体光束平差的延时更低, 而且可以显著地改善无人机视频影像的融合效果。算法使用优化后的相机外方位元素对视频影像做正射纠正, 再把纠正后的影像叠加到三维大场景中。实验结果表明, 这种方法可以在准实时的条件下融合无人机视频影像, 并且相邻视频影像的同名地物点在融合结果中是重合的。

**关键词** 无人机; 视频影像融合; 光束平差; 正射纠正

**中图分类号** P231 **文献标志码** A

## An algorithm for quasi real time fusion of UAV video images

GENG Zhongyuan<sup>1,2</sup>, YUAN Jiying<sup>3</sup>, LI Yingcheng<sup>1,2,4</sup>, SUN Xinbo<sup>1,2</sup>, ZHU Xiang'e<sup>1,2</sup>, FENG Liang<sup>1,2</sup>

(1. Key Laboratory for Aerial Remote Sensing Technology of National Administration of Surveying, Mapping and Geoinformation, Beijing 100039, China; 2. China TopRS Technology Co., Ltd., Beijing 100039, China; 3. Inner Mongolia Land Surveying and Planning Institute, Hohhot 010010, China; 4. Chinese Academy of Surveying and Mapping, Beijing 100830, China)

**Abstract** Unmanned aerial vehicle (UAV) video images can be fused into geographic information system in real time. Homonymous points in adjacent video images are not coincident, because of the errors in the position and attitude data of the UAV. In this paper, a quasi real time fusion method of UAV video images was proposed to solve the problem. Exterior camera elements were optimized with the sliding window bundle adjustment. This method had a lower delay than the bundle adjustment of all video images, and could significantly improve the fusion result of UAV video images. Ortho-rectification of UAV video images was done with the optimized exterior camera elements, and then the result images were overlaid into the 3D large scene. Experimental results showed that UAV video images were fused in quasi real time with this method, and homonymous points in adjacent video images were coincident in the fusion result.

**Key words** UAV; video image fusion; bundle adjustment; ortho-rectification

## 0 引言

无人机技术在近年来飞速发展, 无人机的视频、影像、位置和姿态等数据可以实时地通过无线传输技术回传到地面站, 这些技术拓宽了无人机的应用, 无人机视频领域成为研究的热点之一。文献[1]研究了无人机地理信息视频系统, 实现了视频影像与地理信息平台的实时匹配。文献[2]在视频流上增强了地理信息, 实现了基于视频影像的实时交互, 这项研究使用卡尔曼滤波对无人机的位置和姿态数据做处理, 并结合人工交互来纠正无人机的姿态。文献[3]根据无人机坐标和姿态角计算航拍影像之间的单应变换矩阵, 通过单应变换矩阵的运算完成多幅图像的拼接。上述研究成果虽然使用了无人机的

收稿日期: 2018-07-16

基金项目: 国家重点研发计划子课题项目(2016YFC0803104, 2016YFC0803109)

作者简介: 耿中元(1981—), 男, 辽宁鞍山人, 博士, 主要研究方向为地理信息系统、数据融合以及计算机图形学。E-mail: gzytj@163.com

网络首发时间: 2019-06-24 08:24:02 网络首发地址: <http://kns.cnki.net/kcms/detail/11.4415.P.20190621.1331.016.html>

位置和姿态信息对视频影像做了正射纠正或者增强显示,但是由于飞行过程中气流的干扰以及硬件设备精度的原因,无人机的位置和姿态数据往往存在较大误差,使用这些数据做正射纠正往往会导致相邻影像的同名地物点不能重合。文献[4]根据航拍时的位置姿态信息计算影像之间的变换矩阵,通过最小二乘平差提高了变换矩阵的精度,利用变换矩阵实现无人机多幅影像的拼接。文献[5]提出了局部最优的视频流拼接方法,在拼接时动态选取基准面,优化各影像变换到基准面的绝对单应性矩阵,完成视频流的拼接。文献[6]提出近实时的视频流拼接方法,根据全球定位系统(global positioning system, GPS)数据和给定重叠度构建平衡二叉树,选取离二叉树节点最近的图像作为参考投影面完成拼接。上述3项研究并未优化相机的外方位元素,也没有使用基于共线方程的影像正射纠正方法。经典的GPS辅助空中三角测量<sup>[7-8]</sup>通常用于影像的离线正射纠正,文献[9]分析了曝光延迟对GPS辅助光束法平差的影响,提出了顾及曝光延迟的GPS辅助光束法平差,但是该方法也用于影像的离线正射纠正,不能做到准实时的视频影像融合。在同时定位与建图(simultaneous localization and mapping, SLAM)领域,文献[10-12]研究了快速的图像特征计算与匹配方法以及局部光束平差方法,但是这些方法得到的是相机的相对位置和姿态信息,缺乏地理参考。

为了在准实时条件下实现视频影像融合并解决相邻影像的同名地物点不重合的问题,本文提出了基于滑动窗口的光束平差方法来优化相机的外方位元素,然后对视频影像做正射纠正,再把纠正后的影像叠加到三维大场景中。与全部视频影像做整体光束平差相比,使用滑动窗口做光束平差的延时更低,而且可以显著地改善无人机视频影像的融合效果。经过滑动窗口光束平差处理,在准实时的条件下,相邻的视频影像获得了较好的融合效果。

## 1 融合算法

### 1.1 算法总述

融合算法的流程如图1所示,首先以等时间间隔的方式提取视频影像,使用滑动窗口分批次地对视频影像处理,计算特征点并做特征匹配,然后找到特征点和地物点之间的对应关系。使用前方交会计算出地物点的初始坐标,然后做光束平差,优化相机的外方位元素以及地物点坐标。根据优化后的地物点坐标计算每张影像所覆盖区域的平均高程,然后根据相机的内方位元素和优化后的外方位元素对视频影像做正射纠正,再把正射纠正后的影像叠加到三维大场景中。

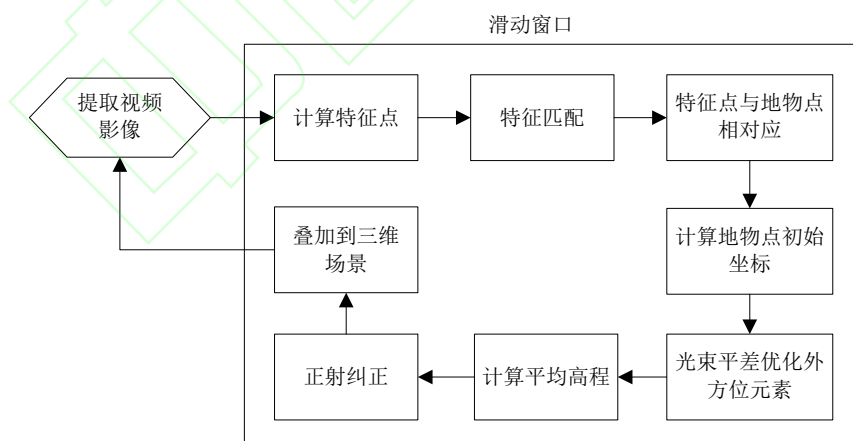


图1 算法流程图  
Fig.1 The Flow Chart of the Method

### 1.2 特征计算与匹配

本文使用有向二进制鲁棒独立基本特征方法(oriented binary robust independent elementary features, 简记为ORB)<sup>[13]</sup>从视频影像中提取并描述特征点,然后使用汉明距离做特征点匹配。ORB方法采用二进制描述特征点,并且有着较好的旋转和缩放不变性,ORB方法的计算和匹配速度非常

快, 适合准实时应用。特征匹配过程中会产生误匹配, 误匹配对平差结果有较大的负面影响需要去除。本文主要解决航拍视频影像的融合问题, 而航拍的视频影像的特征点可以近似认为分布在平面上, 因此本文采用了单应矩阵和随机采样一致性 (random sample consensus, RANSAC) 相结合的方法去除误匹配。对于 2 张视频影像的所有匹配点对, 从中随机选择 4 个点对计算单应矩阵, 再把所有匹配点对代入计算出的单应矩阵, 计算出投影误差在阈值范围内的点对的个数 (即内点的个数), 重复上述过程, 直到内点的个数与总的点对个数的比值大于比例阈值, 或者迭代次数超过预设的最大值。计算出单应矩阵之后把投影误差大于阈值的点对作为误匹配去除。因为融合过程没有考虑视频相机的几何畸变, 所以算法设置了较大的投影误差阈值 (6 个像素), 内点的个数与总的点对个数的比例阈值设置为 0.95, 最大迭代次数设置为 100 次。为了实现准实时的融合, 本文仅对相邻时间间隔的视频影像做两两匹配。

### 1.3 确定特征点与地物点之间的对应关系

平差算法需要根据相机的内方位元素和外方位元素把地物点投影到像平面坐标系, 进而计算出投影坐标, 然后计算投影坐标与特征点坐标之间的误差, 所以需要确定特征点与地物点之间的对应关系。视频影像的重叠度较大, 同名地物点会在多张影像上形成特征点, 算法需要把这些特征点与同名地物点都对应起来, 否则平差算法会把同名地物点当作不同的地物点并分别优化他们的坐标, 这会导致错误的结果。视频影像匹配过程已经生成了匹配点对的列表, 对于第  $k$  张和第  $k+1$  张影像匹配的点对列表, 设  $(x_k, y_k; x_{k+1}, y_{k+1})$  是其中的一个特征点对, 则这两张影像上的特征点  $(x_k, y_k)$  和  $(x_{k+1}, y_{k+1})$  对应于同名地物点, 然后从第  $k+1$  张和第  $k+2$  张影像匹配的点对列表中搜索, 如果能够找到点对  $(x_{k+1}, y_{k+1}; x_{k+2}, y_{k+2})$ , 则说明这 3 张影像上的特征点  $(x_k, y_k)$ 、 $(x_{k+1}, y_{k+1})$  和  $(x_{k+2}, y_{k+2})$  对应于同名地物点。用上述方法遍历所有影像对匹配后生成的点对列表, 找到所有的地物点以及他们对应的特征点。

### 1.4 计算地物点初始坐标

光束平差之前需要计算地物点的初始三维坐标, 本文采用前方交会的方法计算。根据共线方程的反演公式  $\mathbf{p} - \mathbf{p}_s = \lambda \mathbf{R} \mathbf{q}$ , 其中  $\mathbf{p}$  为地物点的物方坐标,  $\mathbf{p}_s$  为相机的物方坐标,  $\mathbf{R}$  为根据相机的外方位元素计算的旋转矩阵,  $\mathbf{q}$  为地物点  $\mathbf{p}$  对应的特征点的像平面坐标,  $\lambda$  为系数, 可以得到地物点的物方坐标  $\mathbf{p}$  的计算公式:  $\mathbf{p} = \lambda \mathbf{R} \mathbf{q} + \mathbf{p}_s$ 。该公式表达的是以相机物方坐标  $\mathbf{p}_s$  为起点, 通过像平面上的特征点  $\mathbf{q}$  的射线, 特征点  $\mathbf{q}$  对应的地物点  $\mathbf{p}$  一定在该射线上。对于每个地物点, 会有两条或者多条以不同的相机坐标为起点, 并且通过相应的像平面特征点的射线, 这些射线的交点即是该地物点的坐标。因为相机的外方位元素以及匹配的特征点的坐标存在误差, 所以这些射线并不会相交于同一个点。本文把这些射线两两求交点, 然后取这些交点的几何中心作为地物点的三维坐标的初始值。实际上空间两条射线不一定会有交点, 因此本文先计算出与两条射线都垂直的线段, 取该垂线段的中点作为两条射线的交点。如果垂线段的端点在射线的反向延长线上, 或者垂线段的端点与相机的距离小于阈值, 则把该端点设置在射线的正方向上, 并且该端点与相机的距离设置为距离的最小阈值, 算法实现时距离最小阈值设置为 1 m。

### 1.5 光束平差

设  $I$  为滑动窗口内所有视频影像的集合,  $\mathbf{P}$  为滑动窗口内所有地物点的集合,  $\mathbf{p}_i$  为任意一个地物点,  $\mathbf{p}_i$  在影像  $I_j$  上的投影记为  $\mathbf{q}'_{ij}$ , 若已知相机的内方位元素和外方位元素以及  $\mathbf{p}_i$  的物方坐标, 则  $\mathbf{q}'_{ij}$  可以使用共线方程计算。设  $\mathbf{q}_{ij}$  为地物点  $\mathbf{p}_i$  在影像  $I_j$  上对应的特征点, 则地物点  $\mathbf{p}_i$  的重投影误差可以写为  $\mathbf{q}'_{ij} - \mathbf{q}_{ij}$ 。可以通过最小化下面的式子来优化滑动窗口内所有相机的外方位元素。



(1)

$$F = \sum_{p_i \in P, I_j \in I} (q'_{ij} - q_{ij})^T \Omega_{ij} (q'_{ij} - q_{ij})$$

式中： $\Omega_{ij}$ 为投影信息矩阵，算法实现时 $\Omega_{ij}$ 设置为单位矩阵。

把式(1)最小化即是在滑动窗口内做光束平差，本文采用通用图优化工具 g2o<sup>[14]</sup>实现光束平差的计算。在优化算法开始之前需要设置优化约束（即图优化算法中的边），在优化约束中实现误差的计算方法，本文采用的误差计算方法即是重投影误差（即 $q'_{ij} - q_{ij}$ ）。优化算法开始之前还需要设置优化变量（即图优化算法中的点），以及优化变量的初始值。本文算法的优化变量分为两类：相机外方位元素和地物点的三维坐标。

设置优化变量之前先要选取坐标系，本文在物方空间坐标实现平差计算，物方空间坐标设置为高斯投影坐标。为了减少高斯投影引起的形变，取第一个相机的 GPS 坐标的经度作为高斯投影的中央经线。高斯投影坐标的数值很大，不利于算法的收敛，因此使用了相对高斯坐标，以第一个相机的高斯投影坐标作为参考点，所有相机的坐标转换为相对于该参考点的坐标。

优化变量中的地物点坐标的初始值采用 1.4 节中描述的方法计算。优化变量中的相机外方位元素的初始值设置方法如下。以第一个相机的高斯投影坐标作为参考点，把滑动窗口内的所有相机的 GPS 坐标转换到相对的高斯投影坐标系下，作为相机外方位的线元素的初始值。因为正射纠正算法使用了以 Y 轴为主轴的转角系统，所以相机的角元素也使用该转角系统，这样平差后的相机外方位元素可以直接用来做正射纠正。但是无人机采用的是航向角、俯仰角、横滚角的转角系统，与正射纠正算法的转角系统并不一致，本文参考文献[15]中的做法，做了转角系统转换，把转换结果作为相机外方位的角元素的初始值。

光束平差的优化过程使用列文伯格-马夸尔特方法，该方法是迭代的过程，每次迭代之后计算相邻两次迭代的误差的比值，如果误差数值接近（算法实现时把该比值设置为 0.995），或者迭代次数超过阈值则迭代过程结束。因为相机外方位元素的初始值可能存在较大误差，所以由前方交会计算出的地物点的初始值也可能存在较大误差。如果优化约束对应的优化变量的初始值的误差较大，则会对优化过程带来负面影响，为了消除这些影响算法做了两次优化。第 1 次优化之后计算所有优化约束的平均误差，然后将平均误差乘以某个误差系数作为误差阈值，把误差大于阈值的优化约束去除，剩余的优化约束和优化变量再做第 2 次优化，算法实现时把误差系数设置为 2.0（即把误差数值大于平均误差 2 倍的优化约束去除）。第 1 次优化过程由于受到误差较大的优化约束的负面影响，优化的结果可能远远偏离了理想的数值，不能在此基础上继续优化，因此，第 2 次优化开始之前重新设置了优化变量的初始值（设置方法与第 1 次优化的初始值设置方法相同）。

### 1.6 正射纠正以及影像叠加

因为视频相机的覆盖范围较小，所以在准实时融合算法中把覆盖范围内的地形近似为平面，正射纠正时用平均高程近似地形。光束平差之后，取优化后的地物点坐标的高程平均值作为每张视频影像覆盖区域的高程的平均值。采用基于共线方程的数字微分纠正方法，使用优化后的相机外方位元素对视频影像做正射纠正，生成带有地理参考的影像。

正射纠正后的影像叠加到中测新图公司的 TopWorld 三维地理信息系统中。为了实现准实时的融合，算法根据视点的地理坐标以及影像的地理范围对正射纠正后的影像做动态的切片，然后通过渲染到纹理（render to texture）的方法，把影像切片渲染到大场景影像纹理中，再把该纹理以贴图的形式贴到大场景地形上，从而实现视频影像与三维大场景的融合。

## 2 算法结果

为了实现准实时的融合，算法实现时使用了多线程。算法流程中从提取视频影像到特征匹配结束这些步骤在每张视频影像获取之后即可执行，所以这些步骤由第 1 个线程完成；从特征点与地物点相

对应到正射纠正这些步骤需要等到窗口内的所有视频影像都获取之后才能执行，所以这些步骤由第 2 个线程完成；每张视频影像在正射纠正之后都可以做影像叠加，所以影像叠加由第 3 个线程完成。

每个滑动窗口包含多张视频影像，如果滑动窗口各自独立地做光束平差，则相邻窗口的交界处的影像之间没有做光束平差，这会导致纠正后的视频影像的同名地物点不重合。为了缓解这个问题，本文在相邻的滑动窗口之间保留了重叠的影像，每个窗口做光束平差时都有重叠的影像作为约束。滑动窗口之间重叠的影像只计算一次特征点，然后做一次特征匹配，匹配的结果在下一个滑动窗口中重复使用。

算法实现时滑动窗口内的视频影像个数设置为 36 张（即每 36 张视频影像做一次光束平差），相邻窗口之间重叠的视频影像个数设置为 6 张，每张视频影像的大小为 1 920 像素×1 080 像素，视频影像的提取时间间隔为 1 s。算法的测试环境是 Windows 7 32 位操作系统，硬件配置为 Intel(R) Xeon(R) W3530 CPU 2.8 GHz、4 GB 内存。通过程序记录了算法步骤的时间消耗，详见表 1。从表 1 中可以看到，从提取视频影像到特征匹配结束，这些步骤处理一张视频影像的平均时间消耗是 0.21 s，处理时间小于视频影像的提取时间间隔，因此不会造成未处理的视频影像的累积。从特征点与地物点相对应到正射纠正这些步骤处理一张视频影像的平均时间消耗是 0.79 s，小于视频影像的提取时间间隔。影像叠加处理一张视频影像的平均时间消耗是 0.15 s，也小于视频影像的提取时间间隔。对于每个滑动窗口而言，窗口内最后一张视频影像的融合延时可以用式（2）计算。

$$T=t_1+nt_2+nt_3+t_4$$

(2)

式中： $t_1$ 为每张视频影像从提取到特征匹配结束的平均时间消耗，在本文的算法结果中 $t_1$ 为 0.21 s； $n$ 是滑动窗口内视频影像的个数，在本文的算法中 $n$ 为 36 张； $t_2$ 是每张视频影像从特征点与地物点相对应到计算平均高程结束的平均时间消耗，在本文的结果中 $t_2$ 为 0.31 s； $t_3$ 是每张视频影像正射纠正的平均时间消耗，在本文的结果中 $t_3$ 为 0.48 s； $t_4$ 是每张视频影像做影像叠加的平均时间消耗，在本文的结果中 $t_4$ 为 0.15 s。所以对于每个滑动窗口而言，窗口内最后一张视频影像的融合延时约为 28.8 s，小于滑动窗口总的影像提取间隔 36 s，所以本文的算法可以实现视频影像的准实时融合。融合延时主要来自光束平差和正射纠正两个步骤，下一步的工作可以从这两方面着手。光束平差之前可以做特征点筛选，减少特征点以减少光束平差的时间消耗。正射纠正可以使用图形处理器（graphics processing unit，GPU）加速，还可以对视频影像做隔片正射纠正，这两种做法都可以降低融合延时。

表1 算法步骤的时间消耗  
Tab.1 Time Costs of the Algorithm Steps

算法步骤	平均时间消耗/(秒 张 <sup>-1</sup> )
从提取视频影像到特征匹配结束	0.21
从特征点与地物点相对应到计算平均高程结束	0.31
正射纠正	0.48
影像叠加	0.15

图 2 是不使用和使用滑动窗口光束平差而生成的影像的对比图，其中图 2（a）的影像是参考文献[15]中的做法，把无人机的位置和姿态做了转角系统转换，然后做正射纠正而生成的，图 2（b）的影像是使用滑动窗口光束平差之后再正射纠正而生成的。图 3 是放大后的影像对比图，从中可以清楚地看到正射纠正前如果不做滑动窗口光束平差，相邻影像的同名地物点会有错位，而使用滑动窗口光束平差之后再正射纠正，相邻影像的同名地物点较好地重合在一起了。

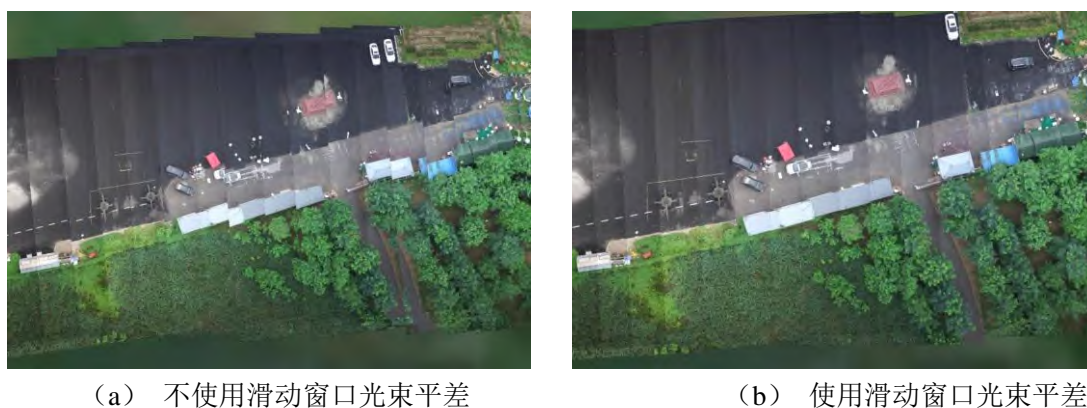


图2 不使用和使用滑动窗口光束平差的影像对比图  
Fig.2 Comparison of Images without and with Sliding Window Bundle Adjustment

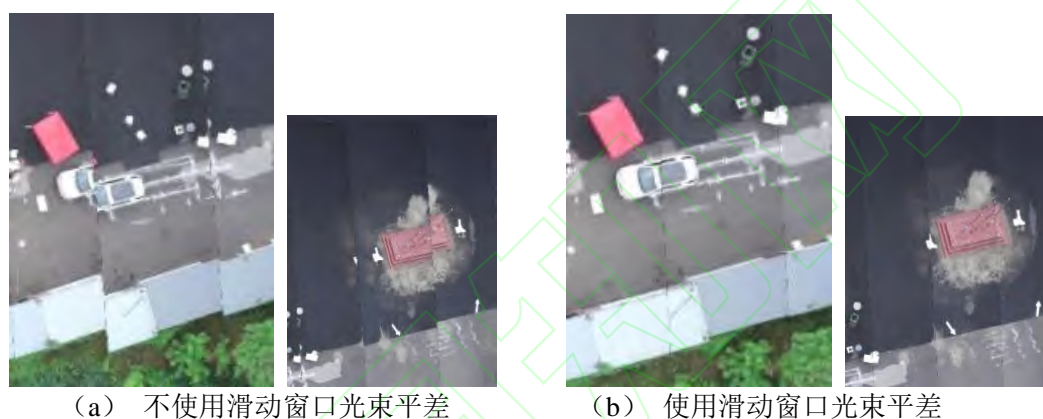


图3 不使用和使用滑动窗口光束平差的放大影像对比图  
Fig.3 Comparison of Scaled Images without and with Sliding Window Bundle Adjustment

本文在相邻的滑动窗口之间保留了重叠的影像，以缓解窗口交界处的同名地物点不重合的问题，这种方法虽然有一定的效果，但是也有副作用。图4和图5是两组对比图，矩形框区域是滑动窗口的交界区域，图4(a)和图5(a)是没有使用重叠影像约束的滑动窗口光束平差的结果图，图4(b)和图5(b)使用了重叠影像约束。针对没有使用重叠影像约束的情况，可以看到在图4(a)和图5(a)中，滑动窗口的交界区域有着明显的错位；而针对使用了重叠影像约束的情况，可以看到在图4(b)和图5(b)中，滑动窗口的交界区域没有明显的错位，同名地物点重合在一起了。在图4(b)的滑动窗口交界的周围区域，同名地物点能够较好地重合在一起，而在图5(b)的椭圆区域，相邻的视频影像之间依然存在着错位的情况。在加入重叠影像约束后，重叠区域的相机的外方位元素是经过优化的（在上一个滑动窗口中做的优化），在重叠区域的边缘，由优化的和未优化的相机外方位元素做前方交计算出的地物点可能存在较大的重投影误差，即使经过光束平差优化，这些区域在正射纠正之后也可能存在同名地物点不重合的情况，这个问题还需要进一步研究。





(a) 不使用重叠影像约束



(b) 使用重叠影像约束

图 4 不使用和使用重叠影像约束的对比图-组 1  
Fig.4 Comparison of Images without and with Overlapped Image Constraint – Group 1



(a) 不使用重叠影像约束



(b) 使用重叠影像约束

图 5 不使用和使用重叠影像约束的对比图-组 2  
Fig.5 Comparison of Images without and with Overlapped Image Constraint – Group 2

### 3 结束语

本文的算法实现了准实时的视频影像融合，通过基于滑动窗口的光束平差方法优化了相机的外方位元素，使用优化后的相机外方位元素对视频影像做正射纠正，使得相邻视频影像的同名地物点较好地重合在一起。本文的算法在正射纠正时使用平均高程近似地形，这种做法适用于地形起伏不大的地区，在地形起伏较大的地区（比如山峰和峡谷），正射纠正过程可以使用 DEM 数据构建三维地形网格并结合 GPU 加速实现快速的正射纠正。本文的算法没有对航带间的影像做光束平差，因为那样做计算量非常大，不能满足准实时的要求，这导致了航带间的影像的同名地物点不能较好地重合。下一步的工作还可以引入 SLAM、计算机视觉等相关领域的方法，以解决航带间的影像的准实时融合问题。



### 参考文献

- [1] 孙新博, 李英成, 王凤, 等. 无人机地理信息视频系统的设计与实现[J]. 测绘科学, 2018, 43(10): 131-136, 156.  
(SUN Xinbo, LI Yingcheng, WANG Feng, et al. Design and implementation of uav geographic information video system[J]. Science of Surveying and Mapping, 2018, 43(10): 131-136, 156.)
- [2] 姜城, 孙敏, 董娜, 等. 面向远程监控的无人机视频地理信息增强方法[J]. 测绘通报, 2014(11): 28-32.  
(JIANG Cheng, SUN Min, DONG Na, et al. Real-time MUAV video augmentation with geo-information for remote monitoring[J]. Science of



- Surveying and Mapping, 2014(11):28-32.)
- [3] 程争刚, 张利. 一种基于无人机位姿信息的航拍图像拼接方法[J]. 测绘学报, 2016, 45 (6): 698-705.  
(CHENG Zhenggang, ZHANG Li. An aerial image mosaic method based on UAV position and attitude information[J]. Acta Geodaetica et Cartographica Sinica, 2016, 45(6): 698-705.)
- [4] 高翔, 朱述龙, 李润生. 位姿信息的小型无人机影像拼接方法[J]. 测绘科学, 2018, 43 (2): 70-76, 82.  
(GAO Xiang, ZHU Shulong, LI Runsheng. A image mosaic method based on position and attitude information for small UAV images[J]. Science of Surveying and Mapping, 2018, 43(2): 70-76, 82.)
- [5] 郑晖, 黎晓东, 渠永平. 无人机视频流倾斜影像快速拼接方法[J]. 科学技术与工程, 2016, 16 (32): 263-268, 282.  
(ZHENG Hui, LI Xiaodong, QU Yongping. Oblique video flow mosaic based on unmanned aerial vehicle[J]. Science Technology and Engineering, 2016, 16(32): 263-268, 282.)
- [6] 郑晖, 姜城, 孙敏, 等. 一种近实时的无人机视频流拼接方法[J]. 地理与地理信息科学, 2015, 31 (5): 1-5.  
(ZHENG Hui, JIANG Cheng, SUN Min, et al. A nearly real-time UAV video frame mosaic method[J]. Geography and Geo-Information Science, 2015, 31(5): 1-5.)
- [7] 袁修孝. POS 辅助光束法区域网平差[J]. 测绘学报, 2008, 37 (3): 342-348.  
(YUAN Xiuxiao. POS-supported bundle block adjustment[J]. Acta Geodaetica et Cartographica Sinica, 2008, 37(3): 342-348.)
- [8] 袁修孝, 朱武, 武军娜, 等. 无地面控制 GPS 辅助光束法区域网平差[J]. 武汉大学学报 (信息科学版), 2004, 29 (10): 852-857.  
(YUAN Xiuxiao, ZHU Wu, WU Junli, et al. GPS-supported bundle block adjustment without ground control points[J]. Geomatics and Information Science of Wuhan University, 2004, 29(10): 852-857.)
- [9] 张春森, 朱师欢, 臧玉府, 等. 顾及曝光延迟的无人机 GPS 辅助光束法平差方法[J]. 测绘学报, 2017, 46 (5): 565-572.  
(ZHANG Chunsen, ZHU Shihuan, ZANG Yufu, et al. GPS-supported bundle adjustment method of UAV by considering exposure delay[J]. Acta Geodaetica et Cartographica Sinica, 2017, 46(5): 565-572.)
- [10] MOURAGNON E, LHUILLIER M, DHOME M, et al. Real-time localization and 3D reconstruction[C]// Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.. Washington, DC, USA: IEEE Computer Society, 2006: 363-370
- [11] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces[C]// Proceedings of the 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, Nara, Japan. Washington, DC, USA: IEEE Computer Society, 2007: 1-10.
- [12] MUR-ARTAL R, MONTIEL J M M, TARDÓS J D. ORB-SLAM: a versatile and accurate monocular SLAM system[J]. IEEE Transactions on Robotics, 2015, 31(5): 1147-1163.
- [13] RUBLEE E, RABAU D V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF[C]// Proceedings of the 2011 IEEE International Conference on Computer Vision (ICCV), Barcelona, Spain. Washington, DC, USA: IEEE Computer Society, 2011: 2564-2571.
- [14] KUMMERLE R, GRISETTI G, STRASDAT H, et al. g2o: a general framework for graph optimization[C]// Proceedings of IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China.[S.l.]:[s.n.], 2011: 3607-3613.
- [15] 袁修孝, 张雪萍, 付建红. 高斯-克吕格投影坐标系下 POS 角元素的转换方法[J]. 测绘学报, 2011, 40 (3): 338-344.  
(YUAN Xiuxiao, ZHANG Xueping, FU Jianhong. Transformation of angular elements obtained via a position and orientation system in Gauss-Kruger projection coordinate system[J]. Acta Geodaetica et Cartographica Sinica, 2011, 40(3): 338-344.)