

近似平面场景多视点图像拼接算法*

徐玉华^{1,2}, 龚文全², 苏昂¹, 张跃强¹, 张小虎¹

(1. 国防科技大学 航天科学与工程学院, 湖南 长沙 410073;

2. 96634 部队, 江西 南昌 330200)

摘要:以小型无人机对地观测为应用背景,研究了近似平面场景多视点图像拼接问题。对于已知粗略相机位姿的情况,提出一种融合相机位姿信息和图像特征点对应信息的方法,采用直接稀疏 Cholesky 分解方法求解拼接全局优化问题。由该方法得到的拼接结果没有全局变形,局部拼接误差也得到了明显的改善。对于相机位姿未知的情况,先采用 structure-from-motion (SFM)方法恢复相机姿态和场景稀疏结构信息,再采用稀疏全局调整方法获得最终的图像变换参数。通过沙盘图像和真实的航拍图像拼接实验验证了算法的有效性。

关键词:图像拼接;稀疏全局调整;三维重建;无人机

中图分类号:TP391 **文献标志码:**A **文章编号:**1001-2486(2014)02-0148-08

Multi-viewpoint image mosaicing algorithm for roughly planar scenes

XU Yuhua^{1,2}, GONG Wenquan², SU Ang¹, ZHANG Yueqiang¹, ZHANG Xiaohu¹

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

2. The Army of 96634, Nanchang 330200, China)

Abstract: Aiming at applications of small unmanned aerial vehicle (UAV) in earth observations, this research explores the multi-viewpoint image mosaicing problems for roughly planar scenes. When coarse camera poses are known, a method for integrating camera poses and feature correspondences is proposed, in which direct sparse Cholesky factorization algorithm is used to solve the global optimization problems of mosaicing. Global distortions do not exist in the obtained mosaics and local mosaic errors are suppressed effectively. When the camera poses are unknown, a structure-from-motion (SFM) system is used to recover the cameras poses and sparse structure of the scene firstly. Then, the sparse global adjustment is used to refine the transformations. The proposed algorithm is validated using sand table images and real aerial images.

Key words: image mosaicing, sparse global adjustment, 3D reconstruction, unmanned aerial vehicle (UAV)

无人机以其机动灵活、可在阴天云下获取光学影像、可低空获取高分辨率影像、可远距离长航时飞行、可在高危地区探测、可在复杂环境下做复杂航线飞行等一系列优点,成为当代卫星遥感和有人驾驶飞机航空遥感的一种有效补充技术手段^[1]。无人机图像拼接在军事侦察、测绘、环境监测、灾情调查等方面有着广泛的应用^[2-5]。

在过去的十几年中,图像拼接一直是图像处理领域中研究的热点之一,到目前为止已出现许多成熟的算法^[7,12]和商业软件^[8]。多数工作都是针对单视点全景图像拼接(single-viewpoint panorama),即相机绕光心旋转的情形。单视点的图像拼接问题及相关的图像融合方法在文献[20]中得到了很好的总结。而在无人机图像拼

接中,相机位置通常是不断变化的,是一种典型的多视点图像拼接,不能直接采用成熟的单视点全景图像拼接算法。

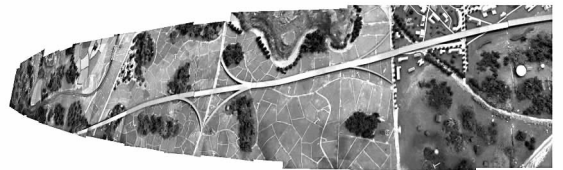


图1 图像配准的累积误差导致拼接结果发生变形(以右边第一张图像作为变换基准)

Fig. 1 Accumulated errors result in distortions of the mosaicing result, (the first image on the right is taken as the reference image)

无人机图像拼接通常需把几百甚至几千张

* 收稿日期:2013-07-08

基金项目:国家973计划项目(2013CB733100);国家自然科学基金资助项目(11272347)

作者简介:徐玉华(1985—),男,江西乐平人,博士,在站博士后,E-mail:robot802@gmail.com;

张小虎(通信作者),男,研究员,博士,博士生导师,E-mail:zxh1302@hotmail.com

图像拼接在一起,形成一张高分辨率拼接图像。如果用帧间图像配准方法^[15],由于存在图像配准误差、非严格平面、镜头畸变等因素,随着图像数量的不断增加,累积误差越来越明显,新加入拼接图的图像可能会发生扭曲^[2,6]。如图1所示,相机以近似正下视的姿态拍摄地形沙盘,在拍摄的过程中,相机的角度和相对于沙盘的高度变化都很小,如果只利用相邻图像之间的变换矩阵对序列图像进行拼接,得到拼接结果虽然看不到明显的拼接缝隙,但图像的左部存在明显的畸变。

1 相关工作

为了避免累积误差的影响,文献[2]用带地理参考信息的卫星图像作为参考图像,把拼接图像与参考图像配准后,叠加在参考图像上,从而生成带地理参考信息的拼接图像^[3]。另一种做法是根据机载GPS/IMU、高度计提供的相机位姿参数以及已知的相机内参,直接对图像进行变换实现图像拼接,被称为基于位姿的方法^[6]。虽然该方法没有累积误差,但由于小型无人机的IMU、GPS、高度计等传感器的精度问题,会导致拼接的图像存在明显的拼接痕迹。为此,文献[6]提出一种混合方法,先采用基于位姿的方法把一幅新的图像加入拼接图中,再用基于图像的方法校正由于相机位置、姿态误差带来的拼接误差,以改善视觉效果。在混合方法中,把每张图像的变换参数 T_i 分成两部分 $T_{i,pos}$ 和 $T_{i,image}$, $T_{i,pos}$ 表示基于相机位姿的变换参数, $T_{i,image}$ 表示基于图像的变换参数。同时调整这两个变换参数,使得定义的目标函数最大化。目标函数定义为^[6]:

$$\lambda(I_{res}) = \alpha \lambda_{spat}(I_{res}) + (1 - \alpha) \lambda_{corr}(I_{res}) \quad (1)$$

$$\lambda_{spat} = \frac{1}{m} \sum_{i=1}^m \frac{1}{\left(1 + \left| \frac{(d_i - \hat{d}_i)}{d_i} \right| \right)}$$

$$\lambda_{corr} = \frac{1}{n} \sum_{i=1}^n \frac{\{1 + CC[Overlaps(I_{res,i-1}, T_i(I_i))]\}}{2}$$

其中 I_{res} 表示拼接的结果图像; λ_{spat} 表示空间距离精度; λ_{corr} 表示在重叠区域图像的相关度,它是图像无缝拼接的度量; $\alpha \in [0, 1]$ 为权重系数,在文献[6]中 $\alpha = 0.5$; CC 表示归一化互相关计算; d_i 表示两个点之间的实测距离, \hat{d}_i 表示根据拼接图像估计的两个点之间的距离。

文献[4]首先用EKF(Extended Kalman Filter)融合GPS和IMU的信息估计相机在各个

时刻的位姿。然后,用EKF估计的结果对匹配的特征点进行三角化计算,得到特征点的三维位置的初始估计。把EKF估计的相机初始位姿和在此基础上得到的特征点的三维位置作为捆集调整^[16,18](bundle adjustment)的初值。最后,根据捆集调整得到的相机位姿和地面特征点的三维坐标把每一幅图像重投影到同一个平面上,形成一幅大尺寸的地面场景拼接图像。文献[14]提出的方法对图像做正射纠正以后再进行拼接,要求有地面控制点和DEM数据。文献[13]只利用单目相机信息,提出一种基于EKF的无人机同时定位与拼接的框架。计算图像之间的单应矩阵时,还计算与之对应的协方差矩阵,作为单应矩阵估计精度的度量。当无人机再次飞越之前已经飞过的区域时(loop-closing),执行EKF流程以减小由累积误差带来的参数漂移,提高图像拼接和无人机的定位精度。文献[15]分析了空中侦察序列图像连续拼接累积误差的原因,为了降低累积误差的影响,提出以中间帧作为最后拼接图像的基准图。这种方法仍不能适用于大量图像的拼接(以20帧图像做测试)。文献[17]提出的结合地理参数的航拍视频实时拼接算法,根据地理参数变换航拍视频,从而简化相机运动模型,使得运动模型只保留水平面平移分量。该方法对传感器的精度有较高要求。

2 本文算法描述

2.1 相机位姿已知的情况

针对配置廉价的GPS/IMU传感器的小型无人机获得的航拍图像,提出一种融合粗略的相机位姿信息和图像配准信息的图像拼接算法。

提出的拼接算法流程如图2所示。首先根据GPS/IMU数据计算相机粗略的外参,对图像进行粗拼接,计算出每幅图像的初始变换参数;然后从图像中提取特征点,在粗拼接的结果指导下进行图像匹配,每幅图像只需要和自身周围的图像进行匹配;为了在图像配准的基础上引入GPS/IMU信息以消除全局累积误差,本文的方法是挑选出部分图像以初始变换参数(由GPS/IMU数据得到)作为最终的变换参数,它们构成了拼接图像的“骨架”,其余的图像的变换参数由全局调整得到。构建“骨架”的目的是防止拼接图像发生整体变形,全局调整的目的是为了改善局部拼接误差。在得到每幅图像的变换参数以后,采用多频

段图像融合 (multi-band blending) 方法^[7] 进行图像融合, 得到最终的拼接结果。

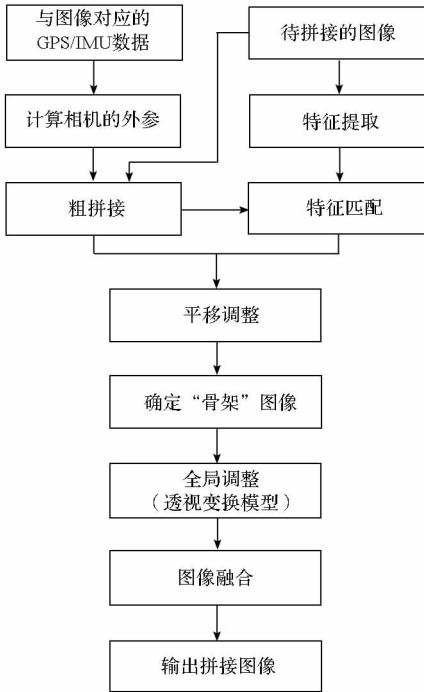


图 2 相机粗略位姿已知情况下的拼接算法流程
Fig. 2 Flow chat of the mosaicing algorithm when coarse camera poses are known

该算法与文献 [6] 提出的方法有相似之处, 主要区别在于文献 [6] 提出的目标函数优化的目标是在 GPS/IMU 提供的位姿约束下的图像区域的相关度最大化, 而本文提出的目标函数的优化目标是在 GPS/IMU 提供的相机位姿约束下的稀疏的匹配图像特征点之间的距离平方和最小化。因此, 本文提出的算法比文献 [6] 具有更小的计算复杂度 (3.1 节给出了相关的比较实验)。另外, 本文采用稀疏 Levenberg-Marquardt (LM) 算法^[16] 进行全局调整, 一方面可以加快优化计算的速度, 另一方面, 使得拼接算法一次能处理更多的图像。

2.1.1 粗拼接

假设已经对相机的内参做过标定。首先用 IMU 输出的无人机姿态和相机安装的姿态以及 GPS 输出的位置信息计算相机的外参 (R, T) 。用相机的内外参数对图像进行下视矫正, 消除滚转角和俯仰角带来的透视畸变; 再根据航向角对图像进行校正, 消除航向角变化引起的图像旋转; 根据各个相机到地面的高度信息将全部图像矫正到相同的相机高度; 再根据 GPS 输出的位置信息得到校正后的图像在拼接图像的坐标。这样就完

成了图像的粗拼接, 从而得到每幅的初始变换参数 (8 自由度的单应矩阵)。

当场景可以假设为一个平面时, 将该平面作为世界坐标系的 $X-Y$ 平面, 图像中的点 $(x\ y\ 1)^T$ (齐次坐标) 与世界坐标系中的平面场景点 $(X\ Y\ 1)^T$ (齐次坐标, $Z=0$) 的映射关系为^[16]

$$\omega(x\ y\ 1)^T = H(X\ Y\ 1)^T \quad (2)$$

其中, ω 为比例因子; H 为 3×3 的单应矩阵, 有 8 个独立的参数, $H = K[r_0\ r_1\ T]$; K 为相机内参矩阵; r_0, r_1 分别为相机旋转矩阵 R 的第 0 列和第 1 列; T 为平移向量。

定义一个标准相机, 它的光轴垂直于世界坐标系 $X-Y$ 平面, 光心坐标和内参矩阵 K 取以所有相机中的某一个为准。设世界坐标系 $X-Y$ 平面到标准相机的图像平面的单应矩阵为 H_s 。以标准相机的图像作为拼接的参考图像。各个相机的图像坐标 $(x\ y\ 1)^T$ 与标准相机的图像 $(x_s\ y_s\ 1)^T$ 之间的关系为:

$$\omega(x_s\ y_s\ 1)^T = H_s H^{-1}(x\ y\ 1)^T \quad (3)$$

利用式 (3) 对各幅图像进行变换, 就完成了图像的粗拼接。

2.1.2 特征提取与匹配

常用的特征提取方法有 Harris^[9]、SIFT^[10]、SURF^[11] 等, 其中 SIFT、SURF 具有旋转和尺度不变性。考虑到 SIFT 在这几种特征中具有最好的鲁棒性, 本文采用 GPU 加速的 SiftGPU^[21] 进行特征提取。

从待匹配的两幅图像 I_A 和 I_B 分别提取 SIFT 特征, SIFT 特征匹配采用基于欧氏距离的最近邻法。对于图像 I_A 中的一个特征点, 利用 $k-d$ 树^[10] 在图像 I_B 的特征中找到欧式距离最近的前两个特征点, 如果最近的距离除以次近的距离少于某个比例阈值, 则接受这一对匹配点。特征匹配完成后, 用 RANSAC 算法^[19] 剔除误匹配对, 如果内点数大于给定的阈值, 则认为图像 I_A 和图像 I_B 匹配成功。

由于在 2.1.1 节中介绍的粗拼接已经确定了每幅图像在总体拼接图中的大致位置, 因此在图像匹配时某一幅图像只需要和空间位置相近的图像进行匹配计算。在图 3 中, 连续的实线表示无人机的航迹, 实心点表示相机拍照的位置。对于在其中的一个由空心点表示的位置处拍摄的图像来说, 它只需要和在虚线框内的位置拍摄的图像做特征匹配计算。

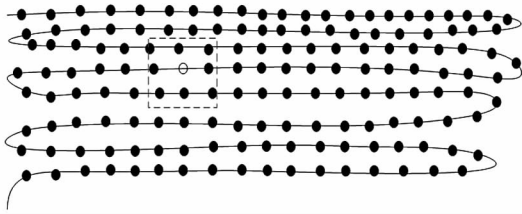


图3 无人机航迹示意图

Fig.3 Flight trajectory of UAV

2.1.3 平移调整

2.1.1 节根据外部传感器信息得到图像的初始变换参数,完成了图像的粗拼接。为了减小由传感器误差带来的图像拼接平移误差,采用2.1.2 节的特征匹配过程建立的图像间的特征点的对应关系对所有的图像同时进行平移调整。

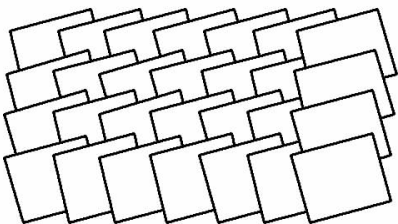
首先用初始变换参数对每幅图像的特征点坐标进行变换;以任意一幅用初始变换参数变换后的图像作为基准,用特征点对应关系建立平移调整方程,用线性最小二乘法求解。用最小二乘法求解线性方程组时,采用2.1.5 节中介绍的直接稀疏 Cholesky 分解方法进行求解。

为了方便叙述,在后文所提到的“粗拼接的结果”,即指对粗拼接进行平移调整后的结果。

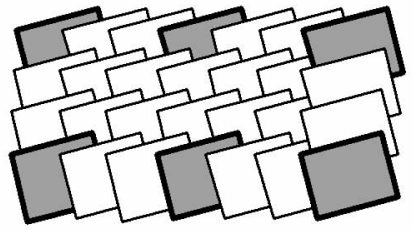
2.1.4 确定“骨架”图像

为了引入 GPS/IMU 的信息以消除图像配准的累积误差,本文提出一种融合 GPS/IMU 信息和图像配准信息的方法。其思路是:根据粗拼接的结果,从待拼接的图像中挑出一部分图像(把它们称为“骨架”图像),这些图像以粗拼接得到的变换参数作为最终的变换参数。而其余的图像的变换参数则以骨架图像为基础,根据图像间的特征点对应关系,通过全局调整计算得到。

假设图4(a)为粗拼接得到的结果,任意挑选一幅图像(本文选第一张图像)作为第一幅骨架图像,其余的骨架图像必须满足条件:骨架图像尽可能的密集,但任意两幅骨架图像的中心点之间的距离至少为图像的长边的 λ 倍(本实验 λ 取1.5)。由这个条件确定的骨架图像如图4(b)所示,粗线框表示骨架图像。



(a) 根据 GPS/IMU 信息进行粗拼接



(b) 确定的骨架图像(粗线框)

图4 确定“骨架”图像

Fig.4 Skeletal images selection

2.1.5 稀疏全局调整

2.1.4 节在粗拼接的基础上挑选出一部分图像构成最终拼接图像的骨架,这些图像以粗拼接得到的图像变换参数作为最终的变换参数,利用这些参数对骨架图像中的特征点坐标进行变换,确定它们在拼接图上的位置。对于其他的图像中的特征点,利用它们之间的对应关系和它们与已经确定在拼接图像上位置的特征点之间的对应关系,建立一个优化目标函数。

假设在全部 M 幅图像之间共找到 N 对特征点, M_f 幅图像的参数已经根据 GPS/IMU 信息确定,其余的 M_u ($M_u = M - M_f$) 幅图像的变换参数设为 X_i, X_i 表示由单应矩阵的8个独立参数组成的列向量,令 $X = (X_1^T, X_2^T, \dots, X_{M_u}^T)^T$ 。目标函数 $f(X)$ 定义为:

$$f(X) = \sum_{i=1}^N e_i^T e_i \quad (4)$$

$$\text{其中, } e_i = \begin{cases} p_{i,m}^f - T_n(p_{i,n}) & i = 1, \dots, N_f \\ T_r(p_{i,r}) - T_s(p_{i,s}) & i = N_f + 1, \dots, N \end{cases}$$

其中, N_f 表示未知变换参数的图像与已知变换参数的图像之间的特征对应个数; $p_{i,m}^f$ 表示利用经过初始变换参数对骨架图像中的特征点变换后的坐标,在式(4)中为常量; $T_n(p_{i,n})$ 表示用第 n 幅图像的变换参数 X_n 对该图像中的一个特征点 $p_{i,n}$ 进行变换的坐标, $(p_{i,m}^f, p_{i,n})$ 为参数未知的图像与参数已知的图像之间的一对匹配的特征点, $(p_{i,r}, p_{i,s})$ 为变换参数未知的两幅图像之间的一对匹配的特征点。

以粗拼接得到的变换参数作为初值,用 LM 算法进行求解。设 E 为单位矩阵, $e = (e_1^T, e_2^T, \dots, e_N^T)^T$, $J = \frac{\partial e}{\partial X}$ 为 Jacobian 矩阵。令

$$A = J^T J + \lambda E$$

LM 算法用以下线性方程迭代求解 X :

$$A \Delta X = -J^T e \quad (5)$$

当图像和特征匹配对的数量不多时,给定初值,用一般的 LM 算法可以很好解决由式(4)表示的优化问题;当待拼接的图像的数量增至几百、几千张时,这时候再用一般的 LM 算法直接求解会非常慢,甚至会因为内存不足导致不可解。

幸运的是,每幅图像中只包含 N 对匹配点对中的部分特征点,因此,式(5)中的 A 是一个稀疏矩阵。本文利用这个稀疏特性,采用 cholmod 计算包^[23]中的直接稀疏 Cholesky 分解方法求解式(5)。

2.2 相机位姿未知的情况

当相机位姿未知时,如有的微型遥控无人机只配置了相机,没有 GPS/IMU 设备。在这种情况下,本文采用 Snavely 等建立的 Structure-From-Motion (SFM) 系统——Bundler^[22]恢复相机的内外参数和场景的稀疏三维点云。根据稀疏的三维点云拟合一个平面,对这个平面、相机参数和三维点云一同进行旋转和平移,使得拟合的平面与世界坐标系的 $X - Y$ 平面重合。这样,就获得了各个视点的相机的参数和一个图像投影平面(世界坐标系 $X - Y$ 平面)。

接下来,再应用 2.1 节中所述的方法进行全局调整和图像融合。由于在 Bundler 的计算过程中,特征提取和匹配已经完成,因此图 2 描述的算法流程中的特征提取、特征匹配就不需要重复进行了。

3 实验结果

3.1 相机位姿已知的情況

为了验证本文算法的有效性,用两组序列图像对算法进行了测试。测试用的计算平台为普通的 PC 机,配置为: Inter Core CPU E7500, 2.93 GHz, 2 GB 内存, NVIDIA GeForce 9600 GT 显卡。

在第一组测试中,以相机运动平台的位置和姿态信息模拟无人机 GPS/IMU 信息。相机以近似正下视的姿态对室内沙盘进行拍摄,拍摄高度约为 1.2 m,有效焦距为 768.3 像素,拍摄的过程中焦距不变,相机基本上沿着一个方向移动,角度和高度变化都很小,共采集了 13 张图像,图像的分辨率为 640×480 ,实验场景如图 5 所示。

图 1 和图 6(a)是基于帧间图像配准的拼接方法得到的结果,其中图 1 以第一张图像作为基准,图 6(a)以中间的图像作为基准(文献[15]的方法),在拼接处它们都具有很好的连续性,没有明显的拼接缝隙,但图 1 左边的部分存在明显的变形,



图 5 室内地形沙盘和 6 自由度相机移动平台
Fig. 5 Indoor terrain sand table and 6-DOF camera moving platform

而图 6(a)通过改变基准图像(以中间的图像作为基准),拼接图像变形有一定程度的改善,但在图像的两端仍然存在由图像配准累积误差等因素带来的不期望的变形。图 6(b)是基于相机位姿信息的拼接结果,相对于图 1 和图 6(a)来说,它没有全局变形,但存在明显的局部拼接误差。图 6(c)是融合相机位姿信息和图像配准信息得到的结果,相比图 6(b)而言,它具有良好的全局一致性,局部的拼接误差也得到了明显的改善。图 6(d)是对图 6(c)用多频段图像融合方法处理后的结果。



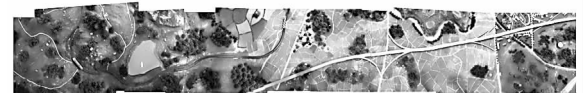
(a) 只用帧间图像配准方法(以中间图像为变换基准)



(b) 只用相机的位姿信息



(c) 稀疏全局调整的结果

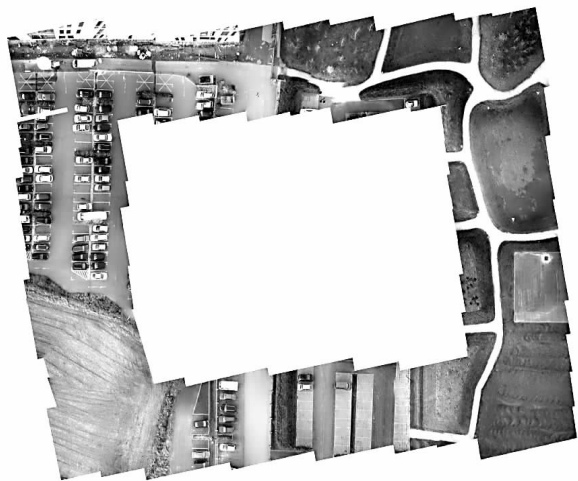


(d) 对(c)进行多频段图像融合处理

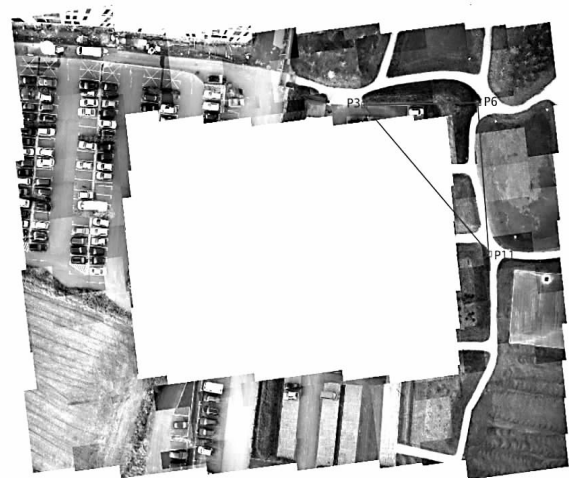
图 6 13 幅沙盘序列图像拼接结果

Fig. 6 Mosaicing results of 13 sequence images of the sand table

在第二组测试中,采用的数据集为文献[6]提供的37幅无人机航拍序列图像(分辨率为 400×300)。由本文算法得到的拼接结果如图7(a)所示(分辨率为 2301×1759),图7(b)为文献[6]算法得到的结果。从视觉效果上看,本文算法得到的结果与文献[6]相当。但本文算法拼接这37幅图像只用了13.9s(CPU 2.93GHz),而文献[6]算法耗费了136s(CPU 2.66 GHz)。本文算法之所以能比文献[6]算法快一个数量级,其主要原因之一在于文献[6]算法目标函数优化的目标是在相机位姿约束下的致密图像重叠区域的相关度最大化,而本文提出的目标函数的优化目标是在相机位姿约束下的稀疏匹配图像特征点之间的距离平方和最小化;另外,本文采用稀疏LM算法求解优化问题也是一个主要因素。



(a) 本文算法结果



(b) 文献[6]算法结果

图7 文献[6]测试图像拼接结果

Fig. 7 Mosaicing result of the test images in references[6]

3.2 相机位姿未知的情况

为了验证2.2节算法的有效性,用手持数码相机对 $7.5\text{m} \times 3\text{m}$ 的沙盘(图5)的各个部分进行拍摄,拍摄时相机的高度、视角、焦距都可以较随意的变化。沙盘中有建筑物、道路、山体、树木、桥梁、河流、草地等模型,山体顶部高度超过 0.3m ,较真实地反映了实际环境的复杂形貌。拍摄过程中,相机到沙盘最低处的高度范围为 $1.2\text{m} \sim 1.5\text{m}$,共采集了111张图像,图像的分辨率为 640×480 ,相机的初始有效焦距从图像的EXIF信息得到。



图8 Bundler从111幅图像得到的沙盘稀疏重建

Fig. 8 Sparse reconstruction of the sand table from 111 images by Bundler

由Bundler得到的沙盘稀疏重建如图8所示。用本文提出的算法得到的沙盘全景图像如图9所示(边缘经过裁剪),两幅拼接图像都采用多频段融合方法进行了处理,其分辨率为 3707×1402 。图9(a)是没有经过全局调整的拼接结果,图9(b)是在Bundler获得相机位姿参数的基础上,再采用2.1节中的全局调整方法得到的结果。经过比较可以看出,经过全局调整后,局部拼接错位得到了明显的改善,如图9(a)下方的道路存在明显的错位,而图9(b)则具有更好的局部一致性。

采用稀疏全局调整的方法拼接111幅图像总共花了 $68\text{min}14\text{s}$,其中Bundler花了 $65\text{min}8\text{s}$ 。如果不采用稀疏的方法,将会因为内存不足而无法进行全局调整计算。

4 总结

对于相机位姿未知的情况,先采用SFM方法恢复相机姿态和场景稀疏结构信息,再采用本文提出的融合相机位姿信息和图像特征点对应信息的方法获得最终图像变换参数。复杂形貌大型沙盘图像拼接实验结果表明,该方法能够得到较满意的视觉效果,但SFM过程耗时较长。在Bundler



(a) 未经过稀疏全局调整的结果



(b) 稀疏全局调整的结果

图 9 由 111 张图像拼成的沙盘全景图像

Fig. 9 The panorama of the sand table from 111 images

的图像匹配过程中,每一幅图像都需要和其他的图像进行匹配。因此,在未来的工作中将利用无人机序列图像的特性,对 Bundler 中的穷举式图像匹配策略进行优化。另外,结构变形处理拼接缝隙的方法^[24]或许能进一步改善本文的拼接结果。

致谢

感谢 Saeed Yahyanejad (Klagenfurt University) 为本文的比较实验提供了测试图像。

参考文献 (References)

[1] 林宗坚,李德仁,胥燕婴.对地观测技术最新进展评述[J].测绘科学,2011,36(4):5-8.
 LIN Zongjian, LI Deren, XU Yanying. General review on the new progress of earth observations[J]. Science of Surveying and Mapping, 2011, 36(4):5-8. (in Chinese)

[2] Heinze N, Esswein M, Krüger W, et al. Automatic image exploitation system for small UAVs[C]. Proceedings of SPIE Airborne Intelligence, Surveillance, Reconnaissance (ISR)

Systems and Applications, 2008.

[3] Se S, Firoozfam P, Goldstein N, et al. Automated UAV-based mapping for airborne reconnaissance and video exploitation[C]// Proceedings of SPIE Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications, 2009.

[4] Bryson M, Reid A, Ramos F, et al. Airborne vision-based mapping and classification of large farmland environments[J]. Journal of Field Robotics, 2010, 27(5):632-655.

[5] 尹鹏飞,尹球,陈兴峰,等.无人机航空遥感技术在震后灾情调查中的应用[J].激光与光电子学进展,2010,47(11):130-134.
 YIN Pengfei, YIN Qiu, CHEN Xingfeng, et al. Unmanned aerial vehicle aerial remote sensing techniques and its application on post earthquake disaster investigation[J]. Laser & Optoelectronics Progress, 2010, 47(11):130-134. (in Chinese)

[6] Yahyanejad S, Wischounig-Struel D, Quaritsch M, et al. Incremental mosaicing of images from autonomous small-scale UAVs[C]//Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance, 2010:329-336.

[7] Brown M, Lowe D G. Automatic panoramic image stitching using invariant features[J]. International Journal of Computer

- Vision, 2007, 74(1):59-73.
- [8] PTGui. <http://www.ptgui.com/>
- [9] Harris C, Stephens M. A combined corner and edge detector[C]//Proceedings of Fourth Alvey Vision Conference, 1988:147-151.
- [10] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91-110.
- [11] Bay H, Tuytelaars T, van Gool L. SURF: Speeded up robust features[J]. European Conference on Computer Vision, 2006:404-417.
- [12] Szeliski R, Shum H Y. Creating full view panoramic image mosaics and environment maps[C]. Proceedings of the 24th annual Conference on Computer Graphics and Interactive Techniques, 1997:251-258.
- [13] Caballero F, Merino L, Ferruz J, et al. Unmanned aerial vehicle localization based on monocular vision and online mosaicking[J]. Journal of Intelligent and Robotic Systems, 2009, 55(4-5):323-343.
- [14] Zhou G Q. Near real-time orthorectification and mosaic of small UAV video flow for time-critical event response[J]. IEEE Transactions on Geoscience and Remote Sensing, 2009, 47(3):739-747.
- [15] 李从利,薛模根,冷晓艳,等.空中侦察序列图像连续拼接的累积误差分析与消除[J].中国图象图形学报,2008,13(4):814-819.
- LI Congli, XUE Mogen, LENG Xiaoyan, et al. Analysis and elimination on aerial recon sequential image stitching accumulative error[J]. Journal of Image and Graphics, 2008, 13(4):814-819. (in Chinese)
- [16] Hartley R, Zisserman A. Multiple view geometry in computer vision[M]. Cambridge:Cambridge University Press, 2003.
- [17] 何贝,王贵锦,沈永玲,等.结合地理参数的航拍视频实时拼接算法[J].应用科学学报,2012,30(2):151-157.
- HE Bei, WANG Guijin, SHEN Yongling, et al. Real-time mosaicing for aerial videos combined with geographical parameters[J]. Journal of Applied Sciences, 2012, 30(2):151-157. (in Chinese)
- [18] Triggs B, McLauchlan P E, Hartley R I, et al. Bundle adjustment—a modern synthesis[C]//International Workshop on Vision Algorithms, Greece, 1999:298-372.
- [19] Fischler M A, Bolles R C. Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography[J]. Communications of the ACM, 1981, 24(6):381-395.
- [20] Szeliski R. Image Alignment and Stitching: A Tutorial[J]. Foundations and Trends in Computer Graphics and Vision, 2006, 2(1):1-104.
- [21] Wu C C. SiftGPU: A GPU implementation of scale invariant feature transform (SIFT) [EB/OL]. <http://cs.unc.edu/~ccwu/siftgpu>, 2007.
- [22] Snavely N, Seitz S M, Szeliski R. Photo tourism: Exploring photo collections in 3D [C]//Proceedings of ACM SIGGRAPH, 2006:835-846.
- [23] CHOLMOD. <http://www.suitesparse.com>
- [24] Jia J, Tang C K. Image stitching using structure deformation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(4):617-631.