引文格式:杨洪,肖腾,吴凌辉,等.面向PTZ相机纯旋转运动的全局式影像定向方法[J].武汉大学学报(信息科学版),2025,50(5):949-958.DOI:10.13203/j.whugis20220712

Citation: YANG Hong, XIAO Teng, WU Linghui, et al. Global Image Orientation Method for PTZ Camera with Pure Rotation [J]. Geomatics and Information Science of Wuhan University, 2025, 50(5):949-958.DOI:10.13203/j.whugis20220712

面向PTZ相机纯旋转运动的全局式影像定向方法

杨 洪[1] 肖 腾[2,3] 吴凌辉[1] 邓 非[1,3]

1 武汉大学测绘学院,湖北 武汉,430079

2 湖北工业大学计算机学院,湖北 武汉,430068

3 武汉天际航信息科技股份有限公司,湖北 武汉,430074

摘 要:Pan-tilt-zoom(PTZ)相机能通过绕光心纯旋转以获得更大视场,被广泛应用于监控系统中。针对在处理PTZ相机影像定向时现有运动恢复结构(structure from motion,SfM)方法存在适用性不足的问题,参考全局式SfM的思路,提出一种面向PTZ相机纯旋转运动的影像定向方法。首先,优选部分像对进行内定向,估计影像的内方位元素;然后,对所有影像进行全局式外定向,估计影像的外方位角元素;最后,利用无物方点的光束法平差优化内方位元素和外方位角元素。仿真数据和真实数据的实验表明,该方法可弥补现有SfM方法的不足,成功实现对PTZ相机纯旋转时的影像定向;在与商用软件对比方面,得到的外定向精度更高,同时具有计算效率上的优势。该方法并不局限于PTZ相机影像的定向,对于多镜头相机、全向相机等设备采集的纯旋转影像的定向问题具有通用性。

关键词:全局式运动恢复结构;影像定向;PTZ相机;纯旋转影像;光束法平差;全向相机

中图分类号:P237 **文献标识码**:A **收稿日期**:2023-04-23

DOI:10.13203/j.whugis20220712 **文章编号**:1671-8860(2025)05-0949-10

Global Image Orientation Method for PTZ Camera with Pure Rotation

YANG Hong[1] *XIAO Teng*[2,3] *WU Linghui*[1] *DENG Fei*[1,3]

1 School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China

2 School of Computer Science, Hubei University of Technology, Wuhan 430068, China

3 Wuhan Tianjihang Information Technology Co. Ltd, Wuhan 430074, China

Abstract: **Objectives**: Pan-tilt-zoom (PTZ) camera is widely used in surveillance systems due to its wide field of view and high resolution. The camera mounted on the gimbal can only make pure rotation motion around the optical center. The optical center overlap of images makes it impossible to use existing structure from motion (SfM) methods for purely rotated images orientation. In order to estimate the elements of interior and exterior orientations in the free net of a set of PTZ camera images with pure rotation, a global image orientation method for PTZ camera with pure rotation was proposed, which motivated by the idea of the global SfM. **Methods**: Compared with the general SfM method, the similarities are feature extraction and feature matching, and the differences are interior orientation, exterior orientation and bundle adjustment. First, partial image pairs are selected for inner orientation to estimate the elements of interior orientation. Then, global exterior orientation was implemented for all images, and the elements of exterior orientation were estimated. Finally, the elements of interior and exterior orientations are optimized by bundle adjustment with no object 3D point. **Results**: Experiment results on simulation data and real data prove the feasibility and accuracy of the proposed method. In the simulation data experiment, the maximum error of the focal length and the principal point is 3.321 pixel, and most of the errors are less than 1 pixel. The maxi-

基金项目:国家自然科学基金(42301491);湖北省重点研发计划(2022BAA035);创新研究专项基金(2022C61540)。

第一作者:杨洪,硕士,主要研究方向为摄影测量。hongyang@whu.edu.cn

通信作者:邓非,博士,教授。fdeng@sgg.whu.edu.cn

mum rotation errors of the four datasets are 0.116°, 0.320°, 0.103°, 0.125°, respectively. In the real data experiment, the maximum reprojection errors of the checkpoints are 4.919 pixel and 4.758 pixel for the two datasets, respectively. **Conclusions:** Compared with the existing global SfM method, the proposed method can successfully orient the images of PTZ camera with pure rotation. In addition, compared with other commercial software, the exterior orientation accuracy of the proposed method is higher, and it also has the advantage of computational efficiency. The method presented is not limited to PTZ camera images, and it has universality for other purely rotated images orientation problems.

Key words: global structure from motion (SfM); image orientation; pan-tilt-zoom (PTZ) camera; pure rotation image; bundle adjustment; omnidirectional camera

Pan-tilt-zoom(PTZ)相机搭载于云台上,能够左右和俯仰旋转以及镜头缩放,具有视场广和清晰度高的优点,被广泛应用于监控系统中[1-2]。为提升PTZ相机的应用价值,基于PTZ相机的全景拼接[3-4]、增强现实[5-6]、目标检测和定位[7-10]得到了广泛的研究。而在这些PTZ影像应用中,影像内外方位元素的精度直接制约了最终的应用效果。为准确估计PTZ影像的内外方位元素,国内外学者进行了研究,主要方法可由是否使用物方参照物进行划分。

文献[11]通过在室内场景布设平面标定板,建立PTZ影像内方位元素与镜头缩放比例的函数关系,以确定影像内方位元素,并借助预生成的场景特征库计算PTZ影像的外方位元素。文献[5]针对运动场的特殊场景,使用其中固定的点线特征计算PTZ影像的内外方位元素。文献[12-13]针对道路场景,使用车道边界线和车辆轮廓线计算影像上的灭点,并通过灭点计算影像的内外方位元素。文献[14]使用普通方形板和二维码作为标定工具,在点的基础上加入直线和平面约束信息,以提高相机标定精度。这些使用参照物的定向方法局限于特定条件和场景,且关注的是单张影像的定向问题,不适合解决一组影像的定向问题。

在不使用物方参照物的定向方法中,文献[15]在给出两张单站PTZ影像外方位元素的情况下,直接分解随机采样一致性算法(random sample consensus,RANSAC)估计的单应矩阵,获得影像内方位元素估计初值,再通过最小化无穷单应约束下的像素几何误差优化内方位元素,但这种方法没有解决外方位元素求解的问题。文献[16]使用人工刺点方式获得一组纯旋转影像间的高精度单应矩阵,再利用单应矩阵列关于绝对二次曲线的方程,解纯旋转影像内方位元素,这种方法仍然没有解决外方位元素估计的问题。文献[17]使用深度学习的方法直接估计两张PTZ影像内方位元素和相对旋转角度,但精度低于传统方法。此外,文献[4]在PTZ相机上加装惯性导航单元(inertial measurement unit, IMU),直接获取影像的外方位角元素,但需考虑IMU设备的成本问题。

单站PTZ相机拍摄的影像通常是光心重合的纯旋转影像,不同于相机既有平移又有旋转的一般情况。相机一般运动时的影像定向,可通过运动恢复结构(structure from motion,SfM)完成。SfM的主流方法为增量式和全局式,两种方法的主要区别在于完成特征提取和匹配后影像外参解算的方法不同。完成特征提取和匹配后,文献[18]的增量式方法首先使用种子像对前方交会出物方点,然后后方交会计算出新影像的外方位元素,通过重复地进行前方交会和后方交会逐次计算出每张影像的外方位元素。由于纯旋转影像光心重合,无法进行前方交会,所以不能使用增量式SfM的方法对纯旋转影像进行定向。文献[19]的全局式方法首先由每个匹配像对的本质矩阵分解出相对旋转,再通过旋转平均估计出每张影像的全局旋转,即外方位角元素,然后计算每张影像的相对平移,最后通过平移平均估计出每张影像的全局平移,得到外方位线元素。在纯旋转情况下,不能再从本质矩阵分解出相对旋转矩阵,因此文献[19]的全局式方法并不能直接用于纯旋转影像定向问题。

影像定向的最后一步往往通过光束法平差(bundle adjustment,BA)对相机内外方位元素和物方点位置进行非线性优化,提升解算精度。由于纯旋转影像共心,无法进行前方交会生成物方点,不能使用如文献[20]中的有物方点的光束法平差方法。文献[21]提出了一种利用对极几何关系,无需前方交会得到物方点,直接由匹配对像素坐标构建极线距离误差方程的无物方点光束法平差方法。

为解决单站PTZ相机固定缩放比例,即保持内方位元素不变,纯旋转拍摄的一组影像定向问

题,本文提出了一种面向PTZ相机纯旋转运动的全局式影像定向方法。不同于一般的全局式SfM算法,在内方位元素解算上,该方法首先优选部分影像进行内定向,在运动平均前估计出准确的内参数。在相对旋转的计算上,全局式SfM从本质矩阵中分解相对旋转和相对平移,本文方法则从单应矩阵中分解出相对旋转。在光束法平差上,对于纯旋转影像无法前方交会出物方点的问题,本文提出了基于无穷单应关系的无物方点光束法平差方法。

1　面向相机纯旋转运动的全局式影像定向方法

本文方法流程图如图1所示,步骤如下:(1)对于一组PTZ相机做纯旋转运动拍摄的影像,本方法基于尺度不变特征变换算法进行特征提取、特征匹配,并根据匹配关系利用AC(a contrario)-RANSAC算法[22]计算单应矩阵。(2)内定向先由像对的单应矩阵解算影像的焦距初值,并设像主点位于影像中心,畸变参数初值为零,再优选匹配点数量最多的部分像对优化焦距、像主点和畸变参数,作为所有影像的内方位元素。(3)外定向使用内定向计算的内方位元素,先由单应矩阵和内参矩阵计算出每个匹配像对的相对旋转矩阵,再通过旋转平均计算所有影像的全局旋转矩阵,即每张影像的外方位角元素。(4)无物方点光束法平差利用无穷单应关系,由所有匹配点对列像素几何误差方程,联合优化内定向和外定向解算的内外方位元素。

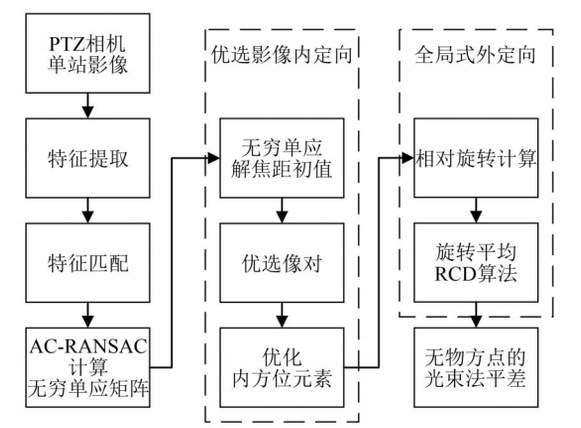


图1　纯旋转影像的全局式定向流程图

Fig. 1　Flowchart of Global Orientation for Images with Pure Rotation

1.1　优选影像内定向

1.1.1　基于无穷单应的焦距初值计算

由于纯旋转影像间没有相对位移,满足无穷单应的条件,焦距初值可通过单应矩阵简便地计算得到。假设两张共视影像 i 和 j 的摄像机矩阵分别为 $\boldsymbol{K}_i[\boldsymbol{E}|0]$ 和 $\boldsymbol{K}_j[\boldsymbol{R}|\boldsymbol{t}]$,其中 $\boldsymbol{K}$ 表示相机内参矩阵,$\boldsymbol{E}$ 和 $\boldsymbol{R}$ 表示相机外参角元素旋转矩阵,0 和 $\boldsymbol{t}$ 表示相机外参线元素。则由空间任意平面 $\boldsymbol{P}=(\boldsymbol{n}^{\mathrm{T}},d)^{\mathrm{T}}$ 诱导的单应为[23]:

$$\boldsymbol{H}=\boldsymbol{K}_j\left(\boldsymbol{R}-\frac{\boldsymbol{t}\boldsymbol{n}^{\mathrm{T}}}{d}\right)\boldsymbol{K}_i^{-1} \tag{1}$$

式中,$\boldsymbol{n}$ 为平面 $\boldsymbol{P}$ 的单位法向量;d 为平面 $\boldsymbol{P}$ 到第一个相机光心的垂直距离。当 d 无穷远时,$\boldsymbol{t}/d$ 趋近于0,即单应是由空间无穷远平面 $\boldsymbol{P}_\infty$ 诱导的无穷单应:

$$\boldsymbol{H}_\infty=\boldsymbol{K}_j\boldsymbol{R}\boldsymbol{K}_i^{-1} \tag{2}$$

设一个相机内参矩阵为 $\begin{bmatrix}\alpha & s & p_x\\ 0 & \beta & p_y\\ 0 & 0 & 1\end{bmatrix}$,其中 α 和 β 分别表示相机在 x 和 y 方向上的焦距;s 表示扭曲参数;p_x 和 p_y 表示像主点。对于一般相机,扭曲参数 $s=0$,且影像宽高比为1,即 $\alpha=\beta$。若只需估计焦距,可假设 $p_x=0$,$p_y=0$,则相机可简化为 $\begin{bmatrix}\alpha & 0 & 0\\ 0 & \alpha & 0\\ 0 & 0 & 1\end{bmatrix}$。此时,无穷远平面 $\boldsymbol{P}_\infty$ 上的绝对二次曲线在该相机内参矩阵下成的像 $\boldsymbol{\omega}$ 可由焦距 α 表示:

$$\boldsymbol{\omega}=(\boldsymbol{K}\boldsymbol{K}^{\mathrm{T}})^{-1}=\begin{bmatrix}1/\alpha^2 & 0 & 0\\ 0 & 1/\alpha^2 & 0\\ 0 & 0 & 1\end{bmatrix} \tag{3}$$

在双视图情况下,绝对二次曲线在影像 i 和 j 下成的像 $\boldsymbol{\omega}_i$ 和 $\boldsymbol{\omega}_j$ 与两张影像间无穷单应 $\boldsymbol{H}_\infty$ 的关系为:

$$\boldsymbol{\omega}_i=\boldsymbol{H}_\infty^{\mathrm{T}}\boldsymbol{\omega}_j\boldsymbol{H}_\infty \tag{4}$$

若已知无穷单应 $\boldsymbol{H}_\infty$,将其代入式(4),可以得到关于 $\boldsymbol{\omega}_i$ 和 $\boldsymbol{\omega}_j$ 中元素的方程:

$$\boldsymbol{A}\boldsymbol{X}=0 \tag{5}$$

以此建立影像内方位元素与无穷单应矩阵的联系。式(5)中,系数矩阵 $\boldsymbol{A}$ 和解向量 $\boldsymbol{X}$ 分别为:

$$\boldsymbol{X}=\begin{bmatrix}\lambda_i/\alpha_i^2 & \lambda_i & \lambda_j/\alpha_j^2 & \lambda_j\end{bmatrix}^{\mathrm{T}} \tag{6}$$

$$\boldsymbol{A}=\begin{bmatrix}h_{11}^2+h_{21}^2 & h_{31}^2 & -1 & 0\\ h_{11}h_{12}+h_{21}h_{22} & h_{31}h_{32} & 0 & 0\\ h_{12}^2+h_{22}^2 & h_{32}^2 & -1 & 0\\ h_{11}h_{13}+h_{21}h_{23} & h_{31}h_{33} & 0 & 0\\ h_{12}h_{13}+h_{22}h_{23} & h_{32}h_{33} & 0 & 0\\ h_{13}^2+h_{23}^2 & h_{33}^2 & 0 & -1\end{bmatrix} \tag{7}$$

h 表示无穷单应矩阵 $\boldsymbol{H}_\infty$ 中的元素[16]。

对于纯旋转影像,由于影像共心,式(1)中 $\boldsymbol{t}$ 为零向量,则利用AC-RANSAC方法估计出无穷

单应矩阵后,就可以根据式(5)列方程组求出纯旋转影像焦距初值。

1.1.2 优选像对优化内方位元素

计算出焦距初值后,为获得更准确的焦距以及其他内方位元素,按匹配点数目最多的原则,从所有匹配对$\varepsilon_{all}=\{(i,j)|i<j\}$优选部分像对$\varepsilon_{part}=\{(i,j)|i<j\}$,通过最小化关于内方位元素的匹配点像素几何误差函数来优化内方位元素。

对纯旋转影像中的任意一组匹配影像(i,j),其第m对特征匹配点坐标向量$\boldsymbol{x}_{i,m}$和$\boldsymbol{x}_{j,m}$(带畸变)之间的映射关系为:

$$\boldsymbol{x}_{j,m}=\underset{i\to j}{\mathrm{proj}}(\boldsymbol{x}_{i,m},\boldsymbol{k})=\boldsymbol{K}_j\boldsymbol{R}_j\boldsymbol{R}_i^{-1}\boldsymbol{K}_i^{-1}\boldsymbol{x}_{i,m}=\boldsymbol{K}_j\boldsymbol{R}_{ij}\boldsymbol{K}_i^{-1}\boldsymbol{x}_{i,m} \tag{8}$$

式中,$\boldsymbol{R}_i$和$\boldsymbol{R}_j$分别表示影像i和j的全局旋转矩阵;$\boldsymbol{R}_{ij}$为相对旋转矩阵;$\boldsymbol{k}$为径向畸变参数向量,包含3个径向畸变参数k_1、k_2、k_3。由于内外方位元素的误差,将i影像上的点$\boldsymbol{x}_{i,m}$映射到j影像上$\underset{i\to j}{\mathrm{proj}}(\boldsymbol{x}_{i,m},\boldsymbol{k})$,两点并不重合,其像素坐标欧氏距离之差,即像素几何误差为:

$$E_{\mathrm{proj}}=d(\boldsymbol{x}_{j,m},\underset{i\to j}{\mathrm{proj}}(\boldsymbol{x}_{i,m},\boldsymbol{k}))^2 \tag{9}$$

对于优选像对ε_{part}中任意一对匹配像对(i,j),有M_{ij}个特征匹配点,按式(9)列所有匹配点的像素几何误差,构成内方位元素的优化方程:

$$(\alpha,p_x,p_y,k_1,k_2,k_3,\{\boldsymbol{R}_{ij}|(i,j)\in\varepsilon_{part}\})=\underset{}{\mathrm{argmin}}\sum_{(i,j)}^{\varepsilon_{part}}\sum_{m=1}^{M_{ij}}d(\boldsymbol{x}_{j,m},\underset{i\to j}{\mathrm{proj}}(\boldsymbol{x}_{i,m},\boldsymbol{k}))^2 \tag{10}$$

式中,$\{\alpha,p_x,p_y,k_1,k_2,k_3\}$为影像组的内方位元素,焦距初值$\alpha$由式(5)解出,假设像主点$p_x$与$p_y$初值为影像中心,畸变参数$k_1$、$k_2$、$k_3$初值为零;优选像对的相对旋转观测初值$\{\boldsymbol{R}_{ij}|(i,j)\in\varepsilon_{part},i<j\}$,可由式(2)得相对旋转矩阵初值计算公式:

$$\boldsymbol{R}_{ij}=\boldsymbol{K}_j^{-1}\boldsymbol{H}_{ij}\boldsymbol{K}_i \tag{11}$$

式中,$\boldsymbol{K}_i$和$\boldsymbol{K}_j$为匹配像对(i,j)的内参矩阵,由内方位元素初值组成;$\boldsymbol{H}_{ij}$为无穷单应矩阵。

1.2 全局式外定向

1.2.1 相对旋转计算

完成内定向计算出影像的内方位元素后,每张影像的内参矩阵已知,根据式(11)可由每对匹配像对的单应矩阵和内参矩阵计算出影像间的相对旋转。然而,由于单应估计存在误差,式(11)所得的相对旋转矩阵并不准确,因此,需将式(11)所得的旋转矩阵当作初始估计,通过最小化匹配点的像素几何误差进行优化:

$$\boldsymbol{R}_{ij}=\underset{\boldsymbol{R}_{ij}\in\mathrm{SO}(3)}{\mathrm{argmin}}\sum_{m=1}^{M_{ij}}d(\boldsymbol{x}_{j,m},\underset{i\to j}{\mathrm{proj}}(\boldsymbol{x}_{i,m},\boldsymbol{k}))^2 \tag{12}$$

式中,$\boldsymbol{R}_{ij}\in\mathrm{SO}(3)$表示相对旋转矩阵在计算时应属于三维旋转群。式(12)不优化内方位元素,将其当作固定值。经过优化后,可得所有影像匹配像对$\varepsilon_{all}=\{(i,j)|i<j\}$的一组相对旋转矩阵观测值$\{\boldsymbol{R}_{ij}|(i,j)\in\varepsilon_{all}\}$。

1.2.2 旋转平均

旋转平均的目的是由相对旋转观测$\{\boldsymbol{R}_{ij}|(i,j)\in\varepsilon_{all}\}$估计出$N$张影像的全局旋转$\{\boldsymbol{R}_i|i\in1,2,\cdots,N\}$,使得影像$i$和$j$间的相对旋转$\boldsymbol{R}_{ij}$与它们的全局旋转$\boldsymbol{R}_j\boldsymbol{R}_i^{\mathrm{T}}$的不一致性最小:

$$\min\sum_{(i,j)}^{\varepsilon_{all}}f(\boldsymbol{R}_j\boldsymbol{R}_i^{\mathrm{T}},\boldsymbol{R}_{ij})^P \tag{13}$$

式中,$f(\cdot,\cdot):\mathrm{SO}(3)\times\mathrm{SO}(3)\mapsto\mathrm{SO}(3)$,表示两个三维旋转群的叉乘在三维空间映射的距离函数,用于度量$\boldsymbol{R}_j\boldsymbol{R}_i^{\mathrm{T}}$和$\boldsymbol{R}_{ij}$的偏差;$P$表示偏差的范数,通常为1或2。为保证旋转平均解的全局性,本文采用旋转坐标下降法,将旋转平均非凸的目标函数(13)通过拉格朗日对偶,转换为凸函数,使得解具有全局最优性[24],即

$$\begin{cases}\min & -\mathrm{tr}(\boldsymbol{Q}\boldsymbol{Y})\\ \text{s.t.} & \boldsymbol{Y}_{ii}=\boldsymbol{E},i=1,2,\cdots,N\\ & \boldsymbol{Y}\succeq0\end{cases} \tag{14}$$

式中,$\boldsymbol{Y}$为全局旋转组成的方阵;$\boldsymbol{Q}$为相对旋转观测值组成的矩阵。

1.3 无物方点光束法平差

由于纯旋转影像共心,无法通过前方交会获得物方点,进行有物方点的光束法平差[20],需采用无物方点的光束法平差方法,它直接利用特征匹配点像素坐标到像素坐标的映射关系,最小化像素几何误差,以优化影像内外方位元素。

式(9)给出了任意一组匹配影像(i,j),由其第m对特征匹配点坐标$\boldsymbol{x}_{i,m}$和$\boldsymbol{x}_{j,m}$(带畸变)像素几何误差的计算公式。对于所有匹配像对$\varepsilon_{all}=\{(i,j)|i<j\}$,每组有$M_{ij}$对匹配点的影像,无物方点光束法平差的误差函数为:

$$E_{\mathrm{BA}}=\sum_{(i,j)}^{\varepsilon_{all}}\sum_{m=1}^{M_{ij}}d(\boldsymbol{x}_{j,m},\underset{i\to j}{\mathrm{proj}}(\boldsymbol{x}_{i,m},\boldsymbol{k}))^2 \tag{15}$$

优化的参数包括每张影像的全局旋转$\{\boldsymbol{R}_1,\boldsymbol{R}_2,\cdots,\boldsymbol{R}_i,\cdots,\boldsymbol{R}_N\}$、一组内方位元素$\{\alpha,p_x,p_y,k_1,k_2,k_3\}$。

2 实验验证与结果分析

为测试本文面向PTZ相机纯旋转运动的全

局式影像定向方法的性能，分别进行仿真数据实验和真实数据实验。仿真数据的采集方法是在场景的三维模型上给出影像的内外方位元素，模拟拍摄过程，生成二维影像。以此方法生成的仿真影像内外方位元素均已知，可以用于计算内外方位元素的真误差，以准确评估解算结果的精度。真实数据利用现实场景中安装的 PTZ 相机采集影像获得，并通过在场景中布置控制点，对本文方法解算结果进行评价。

在 4 个虚拟场景中，采集 4 组仿真影像数据，分别命名为剧院、运动场、城市场景 1 和城市场景 2。在现实场景中，采集了两组真实影像和对应的控制点数据，分别命名为野外场景、城市场景 3。

2.1　旋转精度评价测度

为评价相对和全局旋转精度，定义两个旋转矩阵 $\boldsymbol{R}_1$ 和 $\boldsymbol{R}_2$ 的距离为[25]：

$$l(\boldsymbol{R}_1,\boldsymbol{R}_2)=\arccos((\mathrm{tr}(\boldsymbol{R}_1\boldsymbol{R}_2^{\mathrm{T}})-1)/2) \quad (16)$$

则任意像对 (i,j) 的相对旋转计算精度可表示为 $l(\boldsymbol{R}_{ij}^{c},\boldsymbol{R}_{ij}^{\mathrm{gt}})$，其中，$\boldsymbol{R}_{ij}^{c}$ 为相对旋转计算值；$\boldsymbol{R}_{ij}^{\mathrm{gt}}$ 为相对旋转真值（ground truth，GT）。

由于没有进行绝对定向，全局旋转在一个自由网下，需先将全局旋转的估计值 $\{\boldsymbol{R}_1,\boldsymbol{R}_2,\cdots,\boldsymbol{R}_i,\cdots,\boldsymbol{R}_N\}$ 与真值 $\{\boldsymbol{R}_1^{\mathrm{gt}},\boldsymbol{R}_2^{\mathrm{gt}},\cdots,\boldsymbol{R}_i^{\mathrm{gt}},\cdots,\boldsymbol{R}_N^{\mathrm{gt}}\}$ 对齐，求一个旋转矩阵 $\boldsymbol{R}_{L1}$，使得每组全局旋转距离最小：

$$\boldsymbol{R}_{L1}=\underset{\boldsymbol{R}_{L1}}{\mathrm{argmin}}\sum_{i=1}^{N}d(\boldsymbol{R}_{L1},\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{R}_i^{\mathrm{gt}}) \quad (17)$$

对齐后，再计算对应全局旋转的距离，$l(\boldsymbol{R}_i\boldsymbol{R}_{L1},\boldsymbol{R}_i^{\mathrm{gt}})$ 作为全局旋转精度的评价指标。

2.2　仿真数据实验

仿真数据已知内外方位元素真值，可用于准确评估解算结果精度。首先使用仿真数据评估焦距初值与真值的偏差对外定向的影响，并根据焦距初值计算误差的分布规律选择准确的焦距初值；然后从计算效率和精度两方面考虑，进行优选影像内定向实验，选择合适的优选像对数目；最后在 4 组仿真数据集上进行本文方法与 PhotoScan 的对比实验，比较两种方法的定向性能。

2.2.1　内定向实验

内定向的结果不准确会严重影响外定向精度。采用仿真数据中的城市场景 1 数据进行实验，当焦距与真值偏差越大，相对旋转精度越低，如图 2(a)所示，并进一步影响旋转平均计算的全局旋转精度，如图 2(b)所示。

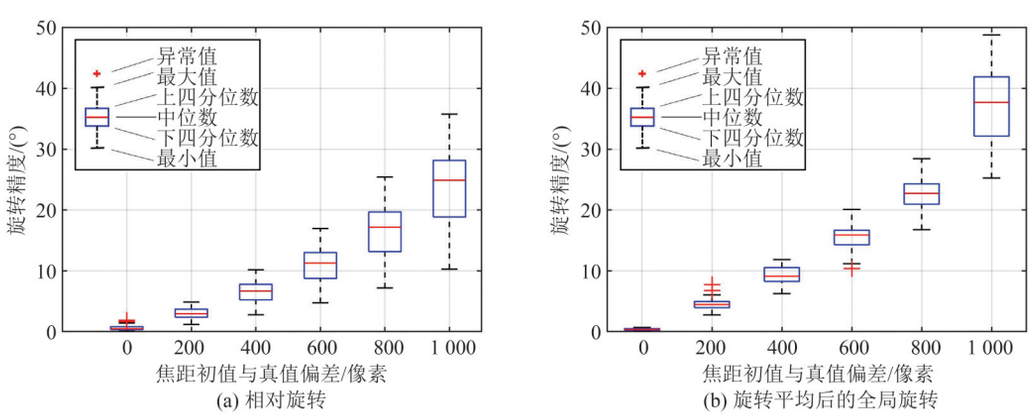


图 2　焦距初值与真值偏差对旋转精度的影响

Fig. 2　Influence of Deviation Between Focal Initial Value and True Value on Rotation Accuracy

焦距误差除降低相对旋转矩阵解算精度外，会进一步使得大量相对旋转观测值在影像三元组几何一致性推论[26]中被剔除，最终导致部分影像的全局旋转估计失败。例如，城市场景 1 数据集共有 66 张影像，在闭合差阈值设置为 5°的情况下，当焦距与真值的偏差达到 200 个像素时，仅能估计出其中 38 张影像的全局旋转，剩余 28 张影像的全局旋转由于缺少观测值而解算失败。因此，在影像外定向前计算准确内方位元素是重要的。

基于无穷单应计算的焦距初值存在误差。原因在于影像特征自动匹配时难以避免地存在较多误匹配点，使得计算的单应矩阵精度不高，致使焦距初值计算不准确。由各场景的所有影像匹配对计算焦距初值，按 100 像素的间隔对所有焦距初值结果进行频次统计，如图 3 所示，大部分解算的焦距初值存在较大误差。

为获得准确的内方位元素，在计算出焦距初值后，§1.1.2 提出了一种优选像对优化内方位元素的方法。由于优选像对内定向的优化函数式(10)为非线性函数，若优选的像对数目过多，则会造成内定向时间过长的问题。为探究像对数量对

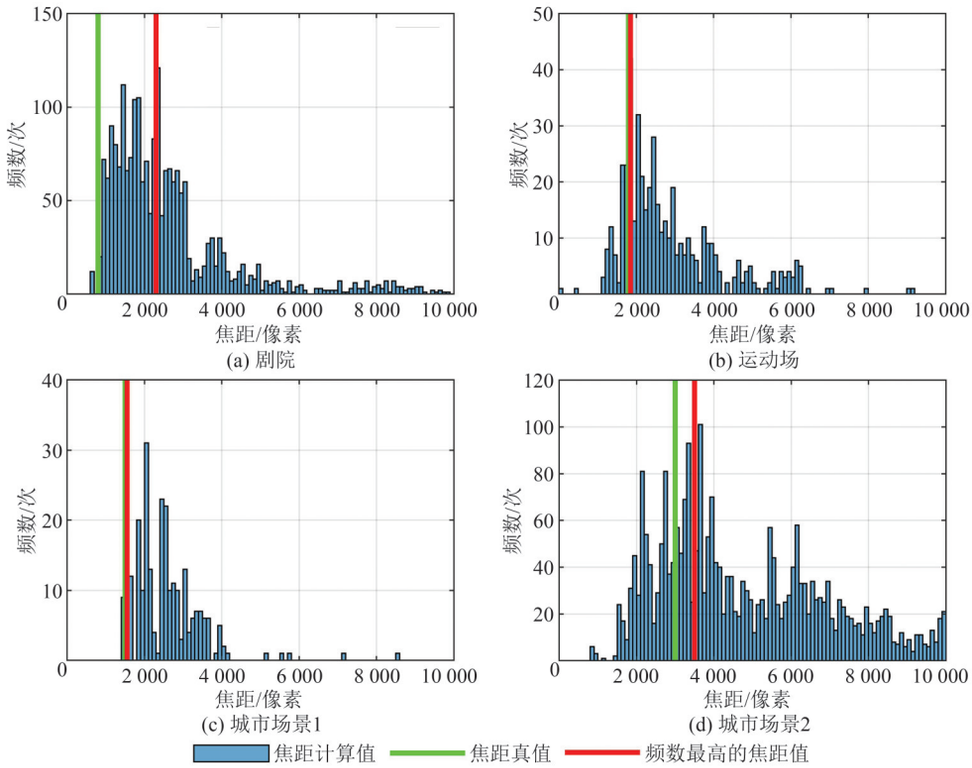


图3 焦距初值频次统计

Fig. 3 Frequency Statistics of Focal Initial Value

内定向精度和时间的影响,对城市场景1数据优选不同数目的像对进行内定向,评估内方位元素的精度和消耗时间。为保证实验准确性,每次优化的内方位元素初值相同,每个指定的像对数目重复50次实验。

当参与内定向的像对数目越多时,标定结果越准确和稳定,如图4所示。但是,像对数目的增大也意味着计算量的增长,从而使内定向更加耗时。综合内方位元素估计精度和时间考虑,本文选择20作为内定向的像对数量。根据实验,当像对数量为20时,焦距值偏差在5个像素内,主点偏差在1个像素内,且耗时小于1 s。

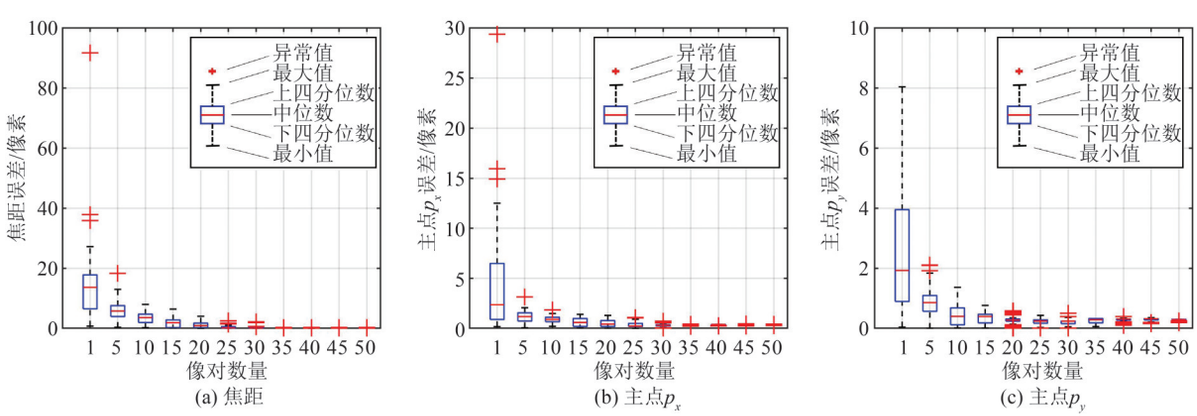


图4 像对数量对内方位元素优化精度的影响

Fig. 4 Influence of the Number of Image Pairs on the Accuracy of Interior Orientation Elements

4组仿真数据内定向的结果如表1所示,经过优选像对优化,此前焦距初值偏差较大的剧院和运动场两数据集,其焦距误差显著降低,为全局式外定向提供了准确的内方位元素。

2.2.2 仿真数据对比实验

对4组仿真数据分别使用本文方法和PhotoScan进行定向,结果对比见表2。在时间耗费方面,本文方法定向流程的耗时,在剧场、运动场和城市场景2这3个数据集中均最短,并且数据集中的影像数目越多,本文方法的计算效率优势越明显。在内方位元素精度方面,本文最终解算的内方位元素与PhotoScan解算的内方位元素精度

相当，畸变参数精度略低于PhotoScan。在旋转精度方面，本文方法的外定向结果优于PhotoScan。

表1　仿真数据优选像对优化内方位元素结果

Table 1　Results of Optimization Within the Selected Image Pairs of Simulation Data

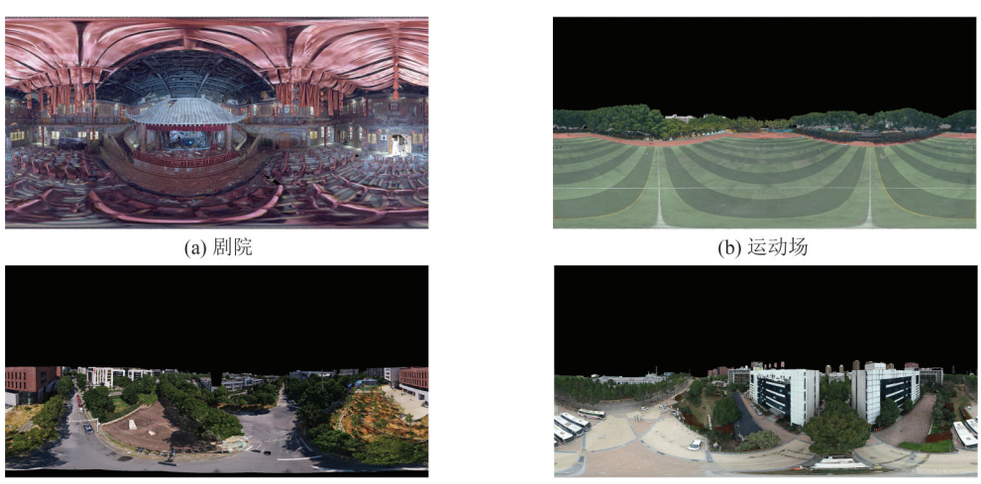
场景	影像数	焦距初值误差/像素	优化后内方位元素误差					
			α/像素	p_x/像素	p_y/像素	k_1	k_2	k_3
剧院	156	1 500.000	0.525	0.813	1.003	0.000	0.048	0.011
运动场	96	0.000	1.310	0.815	1.285	0.007	0.026	0.022
城市场景1	66	0.000	1.830	0.882	0.770	0.023	0.255	0.019
城市场景2	240	600.000	7.640	5.653	4.009	0.025	0.112	0.020

表2　计算时间和方位元素精度对比

Table 2　Comparison of Computation Time and Accuracy of Orientation Elements

数据集	影像数	方法	定向时间/s	内方位元素误差						外方位角元素误差/(°)		
				α/像素	p_x/像素	p_y/像素	k_1	k_2	k_3	最小值	中位数	最大值
剧院	156	本文方法	12	0.493	0.768	0.967	0.002	0.041	0.008	0.090	0.092	0.094
		PhotoScan	22	0.183	0.309	0.414	0.001	0.001	0.001	0.102	0.107	0.114
运动场	96	本文方法	6	0.870	0.598	1.249	0.015	0.154	0.021	0.088	0.104	0.125
		PhotoScan	13	0.880	1.052	0.759	0.010	0.052	0.084	0.099	0.109	0.130
城市场景1	66	本文方法	10	0.410	0.925	0.817	0.023	0.305	0.017	0.082	0.086	0.092
		PhotoScan	8	0.900	1.094	0.248	0.008	0.030	0.020	0.092	0.105	0.119
城市场景2	240	本文方法	26	0.800	3.321	0.876	0.013	0.171	0.020	0.072	0.097	0.122
		PhotoScan	81	1.480	2.326	0.712	0.009	0.039	0.013	0.130	0.178	0.151

定向结束后，直接使用本文方法解算的内外方位元素按圆柱投影[27]生成4组全景影像，如图5所示。4个场景均为PTZ相机监控系统的常用场景，表现出本文方法对不同场景具有较好的通用性。



(a) 剧院　(b) 运动场

(c) 城市场景1　(d) 城市场景2

图5　仿真数据全景拼接结果

Fig. 5　Panoramic Stitching Results of Simulation Data

2.3　真实数据实验

真实数据实验的设备为分别架设于野外和城市高塔的PTZ相机。采集影像时，固定缩放比例，按水平方向30°、竖直方向14°的步长，采集两组环视360°、垂直方向[0°,71°]、每组66张视频帧影像，分别命名为野外和城市场景3数据集。依照本文面向PTZ相机纯旋转运动的全局式影像定向方法进行影像相对定向，得到影像的内方位元素解算结果，并与PhotoScan解算结果进行对比，见表3。根据解算的内外参数拼接的全景影像如图6所示，全景影像无明显错位。

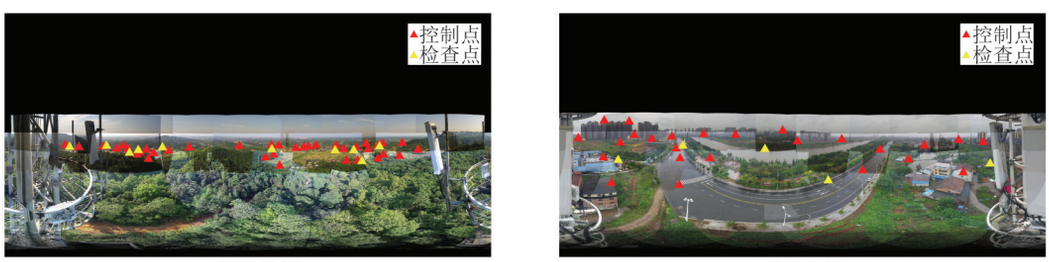
为评价解算结果的准确性，借助控制点为全景影像绝对定向，恢复世界坐标下的位姿，再评估控制点和检查点的重投影误差。具体方法是：

首先在PTZ相机的拍摄范围内均匀布设控制点,测量出控制点的空间坐标;然后在全景影像上刺点,获得控制点像素坐标与世界坐标的对应关系,使用文献[28]的方法进行全景影像绝对定向;最后,将控制点和检查点的世界坐标按绝对定向结果投影到全景影像上,计算重投影误差。由于使用文献[28]的方法进行绝对定向时,不会改变影像间的相对姿态关系,控制点和检查点的重投影误差也会受到全局式外定向的影响,所以重投影误差也可以反映全局式外定向精度。

表3 真实数据定向结果对比

Table 3 Comparison of Orientation Results of Real Data

场景	影像数	方法	步骤	焦距初值/像素	α/像素	p_x/像素	p_y/像素	k_1	k_2	k_3
野外	66	本文方法	内定向结果	3 200	2 815.940	1 292.710	748.349	−0.189	0.042	0.000
			平差后结果	—	2 810.140	1 287.390	745.674	−0.195	−0.088	0.000
		PhotoScan	平差后结果	—	2 807.120	1 287.243	739.822	−0.191	0.162	0.000
城市场景3	66	本文方法	内定向结果	2 400	2 104.780	949.990	568.454	−0.184	0.054	0.000
			平差后结果	—	2 101.130	949.675	572.746	−0.182	0.147	0.000
		PhotoScan	平差后结果	—	2 103.640	958.798	569.203	−0.189	0.154	0.000



(a) 野外场景 (b) 城市场景3

图6 真实数据全景拼接结果及控制点分布

Fig. 6 Panoramic Stitching Result and Control Points Distribution of Real Data

野外场景中,布设47个控制点,选择其中37个用于绝对定向,10个作为绝对定向后的检查点,点位分布见图6(a)。城市场景3中,布设30个控制点,选择其中25个用于绝对定向,5个作为绝对定向后的检查点,点位分布见图6(b)。在两个场景中,先分别用本文方法和PhotoScan进行相对定向,再使用相同方法进行全景影像绝对定向,最后将控制点和检查点的空间坐标按绝对定向结果投影到全景影像上,点位重投影误差见图7。

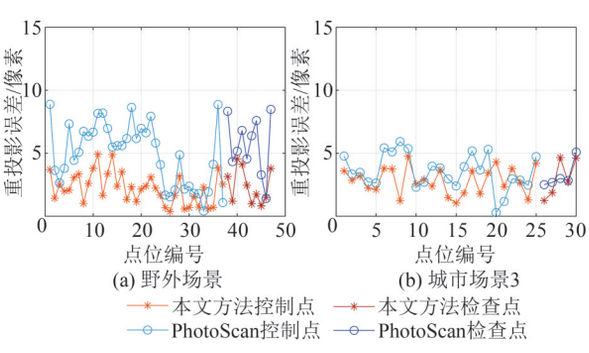


图7 真实数据重投影误差对比

Fig. 7 Comparison of Reprojection Errors of Real Data

在野外场景中和城市场景3中,本文方法和PhotoScan结果控制点和检查点的最大重投影误差分别为4.919、4.546个像素和8.887、8.482个像素。在城市场景3,本文方法和PhotoScan结果控制点和检查点的最大重投影误差分别为4.758、4.618个像素和7.761、8.063个像素。

真实数据实验结果说明,本文的PTZ相机单站影像的全局式定向方法在实际应用中具有可行性,且具有较高的定向精度。

3 结 语

本文将全局式SfM的思想运用于PTZ相机影像定向中,提出了面向PTZ相机纯旋转运动的全局式影像定向方法,该方法主要包括优选影像内定向、全局式外定向和无物方点光束法平差3个部分。优选影像内定向可以求解出PTZ影像准确的内方位元素,提升外定向时相对旋转的计算精度,进而避免大量相对旋转观测值因精度低而被三元组环一致性推论剔除,导致部分影像无法估计出全局旋转的情况,保证影像定向的完整

性。全局式外定向通过从单应矩阵中分解出相对旋转，再进行旋转平均估计出全局旋转，解决了PTZ相机纯旋转影像无法前方交会出物方点，不能使用现有全局式SfM框架定向的问题。无物方点的光束法平差通过最小化匹配点的像素几何误差，联合优化内外方位元素，提升定向结果精度。

仿真数据实验一方面探究了优选像对数量对内定向精度和时间的影响，为内定向选择合适的像对数量提供了参考，对保证定向完整性具有重要意义，另一方面通过真误差评定反映了本文算法的准确性。真实数据实验反映了本文方法实际应用中的可行性，并具有较高的定向精度。鉴于PTZ相机的广泛使用，高精度的PTZ相机定向方法可以在基于城市监控摄像头的城市路网监测[2]、夜光分析[29]、生产实时(准实时)实景三维产品[30]等上层应用中发挥价值。此外，本文PTZ相机单站影像的全局式定向方法也适用于其他纯旋转影像。随着全景视觉在自动驾驶、机器人导航、虚拟现实和增强现实等众多领域的应用[27]，大量的纯旋转摄影设备如泰坦八镜头相机、GoPro全向相机和全景推车被广泛使用，为本文方法提供了广阔的应用空间。

参　考　文　献

[1] SINHA S N. Pan-Tilt-Zoom (PTZ) Camera[M]. Cham: Springer International Publishing, 2021.

[2] 高飞，王美珍，刘学军，等. 一种监控摄像机网络-路网覆盖优化调度方法[J]. 武汉大学学报(信息科学版)，2020，45(3)：362-373.
GAO Fei, WANG Meizhen, LIU Xuejun, et al. A Road Network Coverage Optimization Method in Surveillance Camera Network[J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(3): 362-373.

[3] YONG H W, HUANG J Q, XIANG W M, et al. Panoramic Background Image Generation for PTZ Cameras[J]. *IEEE Transactions on Image Processing*, 2019, 28(7): 3162-3176.

[4] ULLAH F, KWON O J, CHOI S. Generation of a Panorama Compatible with the JPEG 360 International Standard Using a Single PTZ Camera[J]. *Applied Sciences*, 2021, 11(22): 11019.

[5] CHEN J H, ZHU F R, LITTLE J J. A Two-Point Method for PTZ Camera Calibration in Sports[C]//IEEE Winter Conference on Applications of Computer Vision (WACV). Lake Tahoe, NV, USA, 2018.

[6] HOMAYOUNFAR N, FIDLER S, URTASUN R. Sports Field Localization via Deep Structured Models[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017.

[7] 马庆平. 多PTZ主动摄像头的类目标检测定位系统[D]. 成都：电子科技大学，2014.
MA Qingping. Object Detection and Location System Based on PTZ Active Multi-camera[D]. Chengdu: University of Electronic Science and Technology of China, 2014.

[8] YUN K, KIM H, BAE K M, et al. Unsupervised Moving Object Detection Through Background Models for PTZ Camera[C]//The 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 2021.

[9] SIMONSEN C, THIESSON F, HOLTSKOG Ø, et al. Detecting and Locating Boats Using a PTZ Camera with Both Optical and Thermal Sensors[C]//The 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Valletta, Malta, 2020.

[10] 张星，刘涛，孙龙培，等. 一种视觉与惯性协同的室内多行人目标定位方法[J]. 武汉大学学报(信息科学版)，2021，46(5)：672-680.
ZHANG Xing, LIU Tao, SUN Longpei, et al. A Visual-Inertial Collaborative Indoor Localization Method for Multiple Moving Pedestrian Targets[J]. *Geomatics and Information Science of Wuhan University*, 2021, 46(5): 672-680.

[11] WU Z Y, RADKE R J. Keeping a Pan-Tilt-Zoom Camera Calibrated[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(8): 1994-2007.

[12] WANG W, ZHANG Z Y, TANG X Y, et al. Dynamic Self-calibration Algorithm for PTZ Camera in Traffic Scene[C]//Information Technology and Intelligent Transportation Systems, Online Meeting, 2020.

[13] SONG J F, SONG H S, WANG S Y. PTZ Camera Calibration Based on Improved DLT Transformation Model and Vanishing Point Constraints[J]. *Optik*, 2021, 225: 165875.

[14] 谢婧婷，蔺小虎，王甫红，等. 一种点线面约束的激光雷达和相机标定方法[J]. 武汉大学学报(信息科学版)，2021，46(12)：1916-1923.
XIE Jingting, LIN Xiaohu, WANG Fuhong, et al. Extrinsic Calibration Method for LiDAR and Camera with Joint Point-Line-Plane Constraints[J]. *Geomatics and Information Science of Wuhan Universi-*

ty, 2021, 46(12): 1916-1923.

[15] JUNEJO I N, FOROOSH H. Optimizing PTZ Camera Calibration from Two Images[J]. *Machine Vision and Applications*, 2012, 23(2): 375-389.

[16] 郑珂银. 旋转摄像机自标定方法及误差分析[D]. 太原: 山西大学, 2015.
ZHENG Keyin. Self-Calibration Method and Error Analysis of Rotating Camera[D]. Taiyuan: Shanxi University, 2015.

[17] ZHANG C N, RAMEAU F, KIM J, et al. DeepPTZ: Deep Self-calibration for PTZ Cameras[C]//IEEE Winter Conference on Applications of Computer Vision (WACV), Snowmass, CO, USA, 2020.

[18] 陈武, 姜三, 李清泉, 等. 无人机影像增量式运动恢复结构研究进展[J]. 武汉大学学报(信息科学版), 2022, 47(10): 1662-1674.
CHEN Wu, JIANG San, LI Qingquan, et al. Recent Research of Incremental Structure from Motion for Unmanned Aerial Vehicle Images[J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(10): 1662-1674.

[19] MOULON P, MONASSE P, MARLET R. Global Fusion of Relative Motions for Robust, Accurate and Scalable Structure from Motion[C]//IEEE International Conference on Computer Vision, Sydney, NSW, Australia, 2013.

[20] 单杰. 光束法平差简史与概要[J]. 武汉大学学报(信息科学版), 2018, 43(12): 1797-1810.
SHAN Jie. A Brief History and Essentials of Bundle Adjustment[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(12): 1797-1810.

[21] CEFALU A, HAALA N, FRITSCH D. Structureless Bundle Adjustment with Self-Calibration Using Accumulated Constraints[J]. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2016, Ⅲ3: 3-9.

[22] MOULON P, MONASSE P, MARLET R. Adaptive Structure from Motion with a Contrario Model Estimation[C]//Asian Conference on Computer Vision (ACCV), Berlin, Germany, 2013.

[23] 哈特利 R,Hartley Richard,齐瑟曼 A,等. 计算机视觉中的多视图几何[M]. 安徽大学出版社, 2002: 229
Hartley R, Richard H, Zissermana A. Multiple View Geometry in Computer Vision[M]. Hefei: Anhui University Press, 2002: 229.

[24] PARRA Á, CHNG S F, CHIN T J, et al. Rotation Coordinate Descent for Fast Globally Optimal Rotation Averaging[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2021.

[25] LEE S H, CIVERA J. Rotation-Only Bundle Adjustment[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2021.

[26] ZACH C, KLOPSCHITZ M, POLLEFEYS M. Disambiguating Visual Relations Using Loop Constraints[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 2010.

[27] AI H, CAO Z D, ZHU J J, et al. Deep Learning for Omnidirectional Vision: A Survey and New Perspectives[J]. *Computer Vision and Pattern Recognition*, arXiv:2205.1046, 2022.

[28] 邓非, 吴幼丝, 胡玉雷, 等. 球形全景影像位姿估计的改进EPnP算法[J]. 测绘学报, 2016, 45(6): 677-684.
DENG Fei, WU Yousi, HU Yulei, et al. Position and Pose Estimation of Spherical Panoramic Image with Improved EPnP Algorithm[J]. *Acta Geodaetica et Cartographica Sinica*, 2016, 45(6): 677-684.

[29] 陈宇伦, 陈丹妮, 李熙. 利用城市公共摄像头分析夜间灯光变化规律[J]. 武汉大学学报(信息科学版), 2022, 47(8): 1318-1327.
CHEN Yulun, CHEN Danni, LI Xi. Analysis of Night Light Patterns Using Urban Public Camera[J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(8): 1318-1327.

[30] 陈军, 刘建军, 田海波. 实景三维中国建设的基本定位与技术路径[J]. 武汉大学学报(信息科学版), 2022, 47(10): 1568-1575.
CHEN Jun, LIU Jianjun, TIAN Haibo. Basic Directions and Technological Path for Building 3D Realistic Geospatial Scene in China[J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(10): 1568-1575.