

文章编号: 2095-2163(2020)07-0132-05

中图分类号: TP391.4

文献标志码: B

基于深度学习的全景摄像机图像矫正方法研究

张乐江¹, 李晓坤^{1,2}, 张伟华¹, 张懋源¹, 陈虹旭², 杨磊¹

(1 黑龙江恒讯科技有限公司 国家博士后科研工作站, 哈尔滨 150090; 2 黑龙江讯翱科技有限公司, 哈尔滨 150090)

摘要: 本文针对基于深度学习的全景摄像机图像矫正方法进行了综述。阐述了全景相机图像校正和深度学习模型的基本结构,总结了国内外全景相机图像校正技术的研究现状和应用。基于4台400万像素高清鹰眼全景摄像机的全景拼接,通过光学畸变校正对图像校正,建立了现实中真实情境位置与全景相机成像点位置的关系模型,根据实验和实际测得的数据计算出摄像机图形模型参数。建立全景相机图像与规划点图像之间的几何模型,根据摄像机标定模型的图形空间投影变换关系,通过拟合估计矫正参数。根据图像平方拼接算法的原理和过程,建立了虚拟扫描场景的近似等效成像几何模型,生成高精度的远程过程调用参数,实现全景场景拼接校正。

关键词: 深度学习; 图像矫正; 摄像机标定; 投影变换; 全景摄像机

Research on panoramic camera image correction method based on deep learning

ZHANG Lejiang¹, LI Xiaokun^{1,2}, ZHANG Weihua¹, ZHANG Maoyuan¹, CHEN Hongxu², YANG Lei¹

(1 Heilongjiang Hengxun Technology Co., Ltd. Postdoctoral Programme, Harbin 150090, China;

2. Heilongjiang Xun'ao Technology Co. Ltd., Harbin 150090, China)

[Abstract] In this paper, the methods of image correction for panoramic camera based on deep learning are reviewed. The basic structure of panoramic camera image correction and deep learning model are described. The research status and application of panoramic camera image correction technology at home and abroad are summarized. Based on the panoramic Mosaic of 4 4-megapixel HD panoramic cameras and image correction through optical distortion correction, the relationship model between the position of real situation and the imaging point position of panoramic cameras in reality is established, and the camera graphics model parameters are calculated according to the experimental and actual measured data. The geometric model between the panoramic camera image and the planning point image is established, and the correction parameters are estimated by fitting according to the spatial projection transformation relation of the camera calibration model. According to the principle and process of image square stitching algorithm, the approximate equivalent imaging geometric model of virtual scanning scene is established to generate high-precision remote procedure call parameters and realize panoramic scene stitching correction.

[Key words] deep learning; image correction; camera calibration; projection transformation; panoramic camera

0 引言

全国各级公安机关大力开展视频图像建设联网与应用工作,视频图像在建设数量和应用效能上均有显著提升,为公安机关各项业务工作提供了强有力支撑^[1]。然而传统视频图像逐步暴露出“场景小”、“数据散”和“展示平”等问题弊端,全景化展示成为视频图像智慧化发展的趋势^[2],在城市制高点建设鹰眼全景相机,完成180°、270°和360°视角覆盖。而大视角全景图像在实际应用过程中会产生严重的畸变,导致视频图像与实际比例失调,由于摄像

机位置不变、像距保持不变,物距随扫描角度的增大而增大,从而导致图像上从中心到两边比例尺逐渐缩小。垂直高度较高,视角较大,多镜头拼接的非线性因素使畸变复杂化,是常见的一种几何畸变。本文针对这一现实存在的问题,基于深度学习的全景摄像机图像矫正方法研究实现图像的矫正,让图像与肉眼所见的实景更一致。

1 方法概述

1.1 深度学习

深度学习是由人工智能、信号处理、模式识别优

基金项目: 国家自然科学基金(81273649,61501132,61672181);中国博士后科学基金(2019M650069,2020M670939,2019M661319);雏鹰计划(2020CYJBGX0057)。

作者简介: 张乐江(1988-),男,硕士,高级工程师,CCF会员,主要研究方向:物联网、云计算、人工智能等;李晓坤(1979-),男,硕士,研究员级高级工程师,教授,CCF高级会员,主要研究方向:虚拟化、人工智能、生物特征识别等;张伟华(1993-),女,本科生,主要研究方向:智慧医疗、人工智能、生物特征识别等;张懋源(1985-),男,本科生,主要研究方向:人工智能、生物特征识别、智慧城市等;陈虹旭(1986-),男,硕士,工程师,CCF会员,主要研究方向:虚拟化、云计算、人工智能等;杨磊(1989-),男,本科生,主要研究方向:虚拟化、云计算、人工智能等。

通讯作者: 李晓坤 Email:li.xiaokun@163.com

收稿日期: 2020-04-30

化、以及神经网络等多元化学科交叉后创建的新兴研究领域,也可以视作机器学习的一种领域拓展^[3]。

深度学习(Deep Learning)是一种脑神经网络系统,在科学研究的过程中,研究者运用仿生学原理建立框架^[4]。深度学习领域的研究主要是从大量的未标记数据完成对此类未标记数据的特征表示,通过多层结构特征从海量数据中提取简洁、有针对性的数据^[5-6]。该项目也是人工智能、信号处理、模式识别优化等多个学科交叉,通过海量全景相机的图像点与实际的肉眼、计算机视觉和生物信息等之间的缝隙,对人眼进行深度学习,模拟人眼,对图像进行校正^[7]。

1.2 全景摄像机图像矫正

全景摄像机图像矫正是指对失真图像产生的畸变的恢复处理。在实际应用过程中,多个摄像机组合成像系统完成全景拼接造成畸变,从而导致图像失真。本文图像矫正的基本思路是:搭建实际环境;采集所需数据;分析图像失真原因;然后建立相应的数学模型;从摄像机图像已经畸变的图像中,经过分析提取出需要的信息以及相关数据。立足人眼所见的真实环境由图像失真的逆过程恢复图像至现实规划点^[8]。

从两方面入手,一是摄像机定标,二是利用摄像机投影关系,最大程度地还原成真实图像。

2 系统方案与环境搭建

2.1 系统方案

系统主要由以下部分组成:前端设备(摄像机、防雷器、设备箱、支架)、传输网络(光纤收发器或者ONU、光缆)、数据分析与存储。

选取的鹰眼摄像机符合国家相关数据规范的视音频采集。相关设备包含配套支架、设备箱、防雷器、光纤收发器等,电力输入为220 V市电。

传输网络由视频接入网、公安信息网。前端接入部分采用独立组建的视频接入网,完成前端各路视音频的汇聚接入、调用输出和存储转发。其视频信息在视频接入网及通过视频接入网防火墙在公安信息网中应用。

数据分析存储中数据分析以视频的智能应用分析为主。视频智能应用系统以视频图像的智能分析技术为基础,密切贴合实战业务,视频行为智能分析等功能,对重点路段的过车与事件进行监测。

2.2 环境搭建

实地勘察确认安装地点,确保足够立体空间安装全景摄像机、杆件、室外箱等设备。

选择承重墙安装,设计单独的杆件支架,保证安全,要求承重至少为摄像机的10倍以上,承重杆达到300 KG以上。

支架刚性足够,能保证设备在大风情况下不晃动。

摄像机用于高空瞭望大场景,安装高度不低于60 m。

镜头附近不允许出现遮挡物。

180°全景相机安装在楼顶时,相机距离墙面的水平距离建议0.1~1 m。

4台鹰眼全景摄像机的安装相关参数,见表1,当摄像机安装高度为50 m时,其视野盲区(距相机水平距离)为13米。

表1 高度与盲区对应表

Tab. 1 Table of height and blind area

安装高度/m	视野盲区/m
100	24
70	20
50	13
40	11
30	7
20	4
10	2

注:表1中数据为理论计算值,仅供参考,设备实际数据会稍有偏差。

3 数据分析及模型建立

3.1 摄像机标定模型矫正

在全景摄像机镜头的图像畸变中,径向畸变是矢量端点沿长度方向发生矢径的变化 dr 。镜头拼接后的径向畸变在诸多畸变类型中是最为主要的类型^[9]。图1为径向畸变模型。

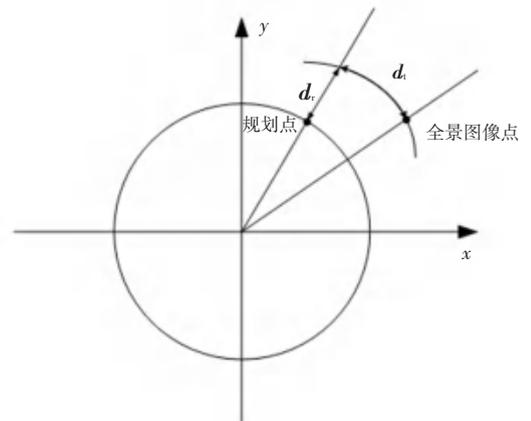


图1 径向畸变模型

Fig. 1 Radial distortion model

根据模型分析全景摄像机镜头的径向畸变,模

型表示为式(1):

$$\begin{aligned} x_{radial} &= x(1 + k_1r^2 + k_2r^4 + k_3r^6), \\ y_{radial} &= y(1 + k_1r^2 + k_2r^4 + k_3r^6). \end{aligned} \quad (1)$$

其中, (x, y) 为全景图像点在图像上的成像坐标, $(x_{tangential}, y_{tangential})$ 为图像矫正后的坐标, 称为矢量端点沿长度方向发生矢径的变化 dr 。实验中普通高清摄像机只计算两项 k_1, k_2 。本次采用的摄像机为单镜头 400 万像素, 总计 1 600 万像素的高精度全景摄像机, 则需要计算 k_1, k_2, k_3 。全景摄像机镜头的切向畸变产生原因, 在于四个镜头系统的组装拼接过程, 全景摄像机镜头的组装一般情况下不能够保证每个镜头的光心与镜头的几何中心都相重合^[10-11]。全景摄像机镜头的畸变数学模型为式(2):

$$\begin{aligned} x_{tangential} &= 2p_1xy + p_2(r^2 + 2x^2), \\ y_{tangential} &= 2p_2xy + p_1(r^2 + 2y^2). \end{aligned} \quad (2)$$

其中, (x, y) 为全景图像点在图像上的成像坐标; $(x_{tangential}, y_{tangential})$ 为图像矫正后的坐标; p_1 与 p_2 为描述全景摄像机摄像头切向畸变程度的待定系数。

3.2 投影变换模型矫正

提取轮廓是提取全景图像中的有效修正区域, 对准备核对矫正区域精准矫正。本文使用扫描线的方法提取全景图像的轮廓, 即从水平方向和垂直方向提取。全景图像是完整的多个扇形组成的区域, 无效区域是黑色。根据研究可知, 每个颜色通道的颜色值是零, 但是从无效区域进入有效区域的过程是阶段性的过程。根据实际试验设置适当的界限是有必要的, 实验中首次超过界限时, 该点就可以被判断为外侧轮廓的边界点之一。对四个方向都进行一次扫描, 可以确定画面的轮廓。接下来根据每个顶点的坐标来确定圆形区域的半径和中心位置。图 2 示意了成像沿某一扫描方向提取轮廓顶点的方法。

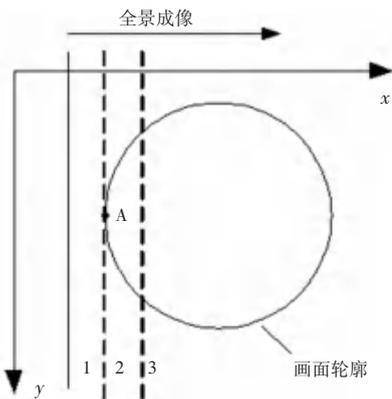


图 2 确定全景成像轮廓示意图

Fig. 2 Determine the outline of panoramic imaging

首先从左到右提取图像, 全景成像线依次通过实线位置 1、虚线位置 2、虚线位置 3。当在虚线位置 2 时, 扫描到一个像素点的位置达到界限临界点, 根据判断这是全景图像轮廓的左顶点。根据得到左顶点的方法依次得到其他三个顶点, 有了 4 个顶点后, 计算轮廓区域的半径和中心坐标。

参数拟合的目的是建立数学模型, 拟合出符合全景透镜特性的参数方程。图 3 是椭球投影变换的正视图。

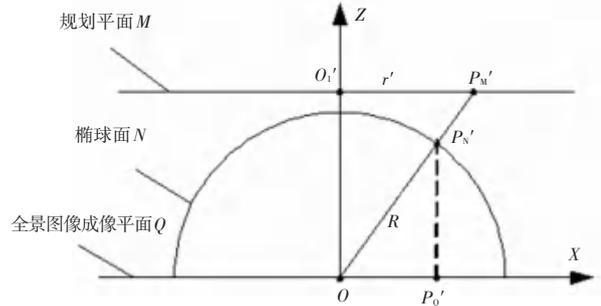


图 3 椭球面投影变换正视图

Fig. 3 Front view of ellipsoid projection transformation

在椭球面投影变换正视图, $P'Q$ 位于成像平面 Q 上; 位于三维立体图像 N 上; 位于全景图像的目标平面 M 上。实验中目标平面与三维立体图像顶点相切, 设置全景轮廓为 R , 那么 $O'_1O' = R, P'_QO' = rO'_1P'_M = r'$, 根据类似三角形的原理, 得到数学关系式(3):

$$\frac{P'_QO'}{O'_1P'_M} = \frac{P'_QP'_N}{O'_1O'} \quad (3)$$

转换得到式(4)

$$\begin{aligned} P'_QP'_N &= \frac{P'_QO'}{O'_1P'_M} \cdot O'_1O' = kR, \\ k &= \frac{r'}{r}. \end{aligned} \quad (4)$$

实验中椭球面上某一点的空间坐标, 与 x, y 存在联系, 对式(2)重新编辑为式(5):

$$(X^2 + Y^2) + k_1Z^2 + k_2(X^2 + Y^2)Z + k_3Z = R^2. \quad (5)$$

全景校正模型基于椭球体, 采样点位于椭球体上。方程(5)中有 K_1, K_2 和 K_3 个未知参数。选择摄像机图像投影变换法拟合待定参数是一种可行的办法。根据实验需求, 选择 3 个采样点来确定相关参数, 根据公式(5)得方程(6):

$$AK = B. \quad (6)$$

转化方程得到方程(7):

$$K = A^{-1}B. \quad (7)$$

其中: A 、 K 、 B 为式(8)

$$A = \begin{bmatrix} \hat{e}_1^2 & (x_1^2 + y_1^2) z_1 & z_1 \\ \hat{e}_2^2 & (x_2^2 + y_2^2) z_2 & z_2 \\ \hat{e}_2^2 & (x_2^2 + y_2^2) z_2 & z_2 \\ \hat{e}_3^2 & (x_3^2 + y_3^2) z_3 & z_3 \end{bmatrix}, \quad (8)$$

$$K = [k_1 \quad k_2 \quad k_3],$$

$$B = \begin{bmatrix} \hat{e}_1 R^2 - (x_1^2 + y_1^2) \\ \hat{e}_2 R^2 - (x_2^2 + y_2^2) \\ \hat{e}_3 R^2 - (x_3^2 + y_3^2) \end{bmatrix}.$$

在实验中,从左、中、右选取 3 个采样点,这样得到的矩阵参数不一定是全局最优值。因此实验进行第二次优化最终利用遍历方法得到全局最优参数数组。因为拟合曲线需要经过中心点,所以选择中心点作为拟合的采样点的第一个落脚点,之后在中心点的两侧各选择一个采样点得到 K 矩阵参数,得到 K_1 、 K_2 、 K_3 、 \dots 、 K_N 。以中心点为基准左侧采样点有 n_1 个,右侧采样点有 n_2 个,计算可得矩阵参数组的个数为 $N = n_1 \times n_2$ 。采样点的理论值 Z 理论可由式(4)计算,采样点的拟合值 Z 拟合可由式(4)计算至式(8)。对式(9)的矩阵参数进行分析,选择满足数据,计算得全局最优值的数据。遍历法具有较低复杂度且能提高算法的稳定性和可靠性,让结果更趋于真实。

$$F(Ki) = \min \left\{ \sum_{i=1}^{n=N} |Z \text{ 拟合 } i - Z \text{ 理论 } i| \right\}, \quad (9)$$

其中, $i = 1, 2, \dots, N, N = n_1 + n_2 + 1$ 。

4 实验验证与分析

4.1 数据采集

现场环境搭建好后,笔记本电脑(配置:第十代英特尔酷睿 i5-10210U 处理器, 8G 内存, 512G SSD 硬盘, MX250 2G 独立显卡)通过六类双绞线连接至设备箱内交换机,通过网络连接至摄像机,并完成对摄像机的相关网络配置。笔记本电脑 WEB 页面登录摄像机,通过自带插件采集视频图像,如图 4 所示。

4.2 实验验证

本次待标定的摄像机型号为海康威视 iDS-2VDP12,采用 4 个 1/1.8" 4MP Progressive Scan CMOS 镜头拼接成全景,设定摄影参数见表 2。校准板的尺寸为 1 920 mm×1 200 mm,由 96 行 60 列,长宽为 20 毫米的正方形组成。对于摄像机采集的图

像,姿态越大,标定结果就越精确。根据运算效率,本次实验定在 150 组到 250 组之间。将摄像机标定图像参数加载到标定系统,计算出全景摄像机的固有参数和畸变系数。表 3 为鹰眼全景摄像机标定结果。



图 4 采集视频图像

Fig. 4 Capture video image

表 2 摄像机 iDS-2VDP12 参数

Tab. 2 Camera iDS-2VDP12 parameters

传感器类型	4 个 1/1.8" 4MP Progressive Scan CMOS
镜头	2.8mm/F1.6
视场角	水平 180°,垂直 87°
视频压缩	H.265/H.264/MJPEG, 支持 Baseline/ Main/ High Profile

表 3 鹰眼全景摄像机标定结果

Tab. 3 Calibration results of eagle eye panoramic camera

项目	类别	参数	结果
标定结果	内参数	x 方向等效焦距 f_x	3 768.605 712 891
		y 方向等效焦距 f_y	3 799.682 434 082
	畸变系数	x 方向主点位置 c_x	3 901.976 867 676
		y 方向主点位置 c_y	2 182.894 042 969
	畸变系数	径向畸变参数 k_1	- 0.257 663 810 849
		径向畸变参数 k_2	0.056 768 815 636 60
		径向畸变参数 k_3	- 0.014 270 393 246 8
		切向畸变参数 p_1	0.003 670 708 892 87
		切向畸变参数 p_2	0.000 122 295 940 45

通过摄像机标定得到的内部参数矩阵和畸变系数,利用 OpenCV 提供的 cvUndistort(2) 函数得到校正后的图像。

假设 P 点位于椭球面上,由于全景成像平面和目标平面处于不同的空间坐标系中,沿点的经度分别圈出 -90° 和 90° 得到 A 点和 B 点,利用坐标系变换来实现投影变换,连接 PA 、 PB 、 OP , 得到 3 个向量 \vec{PA} 、 \vec{PB} 、 \vec{OP} , 其中 \vec{PA} 、 \vec{PB} 满足正交关系,图 5 是摄像机标定图像的重映射误差畸变影像处理效果分析。

通过右手螺旋关系得到式(10):

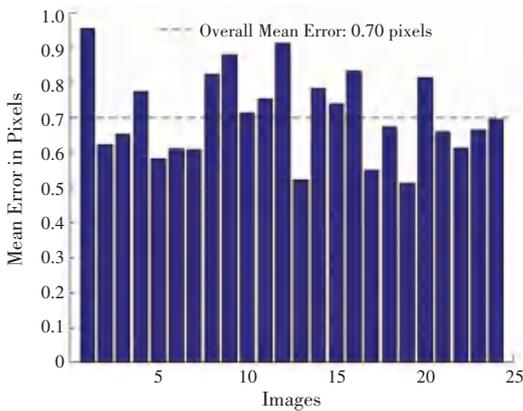


图5 畸变影像处理效果分析

Fig. 5 Analysis of distortion image processing effect

$$\begin{cases} \vec{\alpha} = \vec{PA} \times \vec{PB}, \\ \vec{\beta} = \vec{\alpha} \times \vec{OP}. \end{cases} \quad (10)$$

基于摄像机投影变换实验,得到了一组新的空间坐标系。垂直的俯仰角、水平的旋转角进行空间位置变换,基于深度学习的方法矫正后,可以得到全景校正图像^[12]。通过改变目标平面与实验模型的视距比,还可以得到多视点校正图像,而且是不同变焦度的。椭球投影坐标系 XOZ 有个点 (P_x, P_y, P_z) , 这个点投影到一个规划矫正后的空间坐标系,然后转换为一个新的点 (q_x, q_y, q_z) 。在新的坐标系中,需要对 q_x, q_y 处的 Z 值标准化,以避免由于标准化问题而得到无效图像。对式(5)修改,得式(11):

$$\begin{cases} r^2 \frac{z^2}{q_z^2} + k_1 z^2 + k_2 \frac{z}{q_z} + k_3 z = R^2, \\ r = \sqrt{q_x^2 + q_y^2}. \end{cases} \quad (11)$$

整理得式(12):

$$\begin{cases} (A^2 + k_2 A + k_1) z^2 + k_3 z = R^2, \\ A = \sqrt{\frac{q_x^2 + q_y^2}{q_z^2}}. \end{cases} \quad (12)$$

在椭球面上,计算得到 $Z = Z_0$ 。

在椭球体上, $Z = Z_0$ 。

假设全景轮廓的中心为 (r_x, r_y) , 则目标平面和全景图像平面之间的坐标对应关系可以通过组合式(4)和式(12)得到式(13):

$$\begin{cases} x_{ij} = k_{ij} q_{x-ij} + r_x, \\ y_{ij} = k_{ij} q_{y-ij} + r_x. \end{cases} \quad (13)$$

4.3 实验分析

在原始图像中,假设水平方向上有 N_1 个像素,垂直方向有 N_2 个像素,则 $1 \leq i \leq N_1, 1 \leq j \leq N_2$ 。根据实验模型得到相应的坐标点,用摄像机标定和

投影变换两种模型,基于深度学习的矫正代替坐标 (x_{ij}, y_{ij}) , 完成对全景图像描述,得到矫正后的图像。对采集全景摄像机图像内的道路、河流等明显产生畸变的坐标进行了较好的矫正,使图像趋于真实肉眼所见场景。

5 结束语

本文讨论在城市制高点全景摄像机的实际应用场景中,基于深度学习对全景摄像机图像矫正,利用摄像机标定技术和摄像机投影变换方法,有效地解决了全景相机镜头拼接以后带来的图像几何畸变,恢复到图像的正常视觉比例。本文的算法已在多个摄像机上进行了测试,具有通用性和实用性。在实现图像矫正后,利用高点来级联低点监控、人脸识别、警务通和警车等相关数据,实现在大场景下统一指挥调度,可以为公安工作提供技术支持,有助于安全城市的建设。

参考文献

- [1] 姚天宇. "天网工程"视频监控设备管理存在的问题及对策探讨[J]. 科技经济导刊, 2020, 28(7):16-17.
- [2] 王诗涵. 5G背景下"天网工程"在资金监控方面的运用探究[J]. 财经界, 2020, 28(7):64-65.
- [3] 陈虹旭, 李晓坤. 基于深度学习的指纹识别方法研究[J]. 智能计算机与应用, 2018, 8(3):64-69.
- [4] Barone S, Neri P, Paoli A, et al. 3D acquisition and stereo-camera calibration by active devices: A unique structured light encoding framework[J]. Optics and Lasers in Engineering, 2020, 127: 105989.1-105989.12.
- [5] KANRAR S, MANDAL N K. Video traffic analytics for large scale surveillance[J]. Multimedia Tools and Applications, 2017, 76(11):13315-13342.
- [6] LI W, ZHANG F, ZHANG Y, et al. Adaptive Sample Weight for Machine Learning Computer Vision Algorithms in V2X Systems[J]. IEEE Access, 2019, 7:4676-4687.
- [7] QIANQIAN Z, WEIMING G, YING S, et al. Research on Intelligent Vehicle Damage Assessment System Based on Computer Vision[J]. Journal of Physics: Conference Series, 2020, 1518(1):012050.
- [8] Junhua, Sun, Xiaoqi, et al. Camera calibration based on two-cylinder target[J]. Optics express, 2019, 27(20):29319-29331.
- [9] XU X, YANG H. Vision Measurement of Tunnel Structures with Robust Modelling and Deep Learning Algorithms[J]. Sensors, 2020, 20(17):4945.
- [10] ABDULLAH S, TAHAR K N, RASHID M F A, et al. Camera calibration performance on different non-metric cameras[J]. Pertanika Journal of Ence and Technology, 2019, 27(3):1397-1406.
- [11] CHEN X, FAN R, WU J, et al. Fourier-transform-based two-stage camera calibration method with simple periodical pattern[J]. Optics and Lasers in Engineering, 2020, 133:106121.
- [12] Yingdong, PI, Baorong, et al. On-orbit geometric calibration of linear push-broom optical satellite based on sparse GCPS[J]. Journal of Geodesy and Geoinformation Ence, 2020, 3(1):66-77.