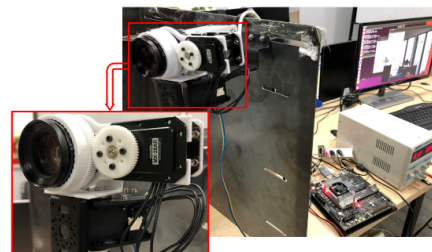




DOI: 10.12086/oe.2020.190436

多场景下基于快速相机标定的柱面图像拼接方法

傅子秋¹, 张晓龙², 余成³, 梁丹¹, 梁冬泰^{1*}¹ 宁波大学机械工程与力学学院, 浙江 宁波 315211;² 上海浦江桥隧运营管理有限公司, 上海 200023;³ 宁波詮航机械科技有限公司, 浙江 宁波 315200

摘要: 针对目前利用相机标定参数进行图像拼接的方法存在受场景限制大、标定过程复杂而耗时长的问题, 提出一种多场景下基于快速相机标定的柱面图像拼接方法。首先, 利用棋盘格标定板角点特征提取精度高的特点, 使其分别位于两两邻接图像的重叠视场中, 对该图像序列依次进行角点提取、精确化和匹配等预处理, 以准确快速求解出待拼接图像间的配准参数; 然后利用标定得到的配准参数快速拼接图像, 通过柱面投影以保持图像的视觉一致性, 并采用多频段融合以保留图像的细节信息; 最后, 将整个系统搭建在低功耗嵌入式平台, 实现可在多场景下完成快速标定及基于标定参数的拼接过程。实验结果表明, 该文方法在室内及隧道等场景下可准确快速完成相机标定, 图像拼接过程耗时短, 同时可保证较高的拼接精度和较好的成像效果, 具有较强的鲁棒性。

关键词: 图像拼接; 相机标定; 柱面投影; 多频段融合

中图分类号: TP317.4

文献标志码: A

引用格式: 傅子秋, 张晓龙, 余成, 等. 多场景下基于快速相机标定的柱面图像拼接方法[J]. 光电工程, 2020, 47(4): 190436

Cylindrical image mosaic method based on fast camera calibration in multi-scene

Fu Ziqiu¹, Zhang Xiaolong², Yu Chen³, Liang Dan¹, Liang Dongtai^{1*}¹ Faculty of Mechanical Engineering and Mechanics, Ningbo University, Ningbo, Zhejiang 315211, China;² Shanghai Pujiang Bridge and Tunnel Operation Management Co., Ltd, Shanghai 200023, China;³ Ningbo Hermeneutic Aircraft Machinery Technology Co., Ltd, Ningbo, Zhejiang 315200, China

Abstract: A cylindrical image mosaic method based on fast camera calibration in multi-scene is proposed to solve the problems of scene limitation and complex calibration process in image mosaic using camera calibration parameter. Firstly, the accurate corner feature of checkerboard calibration board is used to make it in the overlapping field of view of two adjacent images. Then, the image sequence is pre-processed by corner extraction, precision and matching, so that the registration parameters between the images to be stitched can be solved accurately and

收稿日期: 2019-07-24; 收到修改稿日期: 2019-11-29

基金项目: 国家自然科学基金青年基金资助项目(51805280); 浙江省公益性技术应用研究计划(2017C31094); 浙江省自然科学基金资助项目(LQ18E050005)

作者简介: 傅子秋(1995-), 男, 硕士, 主要从事机器人感知与操控、机器视觉图像处理的研究。E-mail: fuziqiuno1@163.com

通信作者: 梁冬泰(1981-), 男, 博士, 副教授, 主要从事工业视觉检测、机器人感知及操控技术的研究。E-mail: liangdongtai@nbu.edu.cn

版权所有©2020 中国科学院光电技术研究所

quickly. After that, the cylindrical projection is used to maintain the visual consistency of the images, and the multi-band fusion is used to retain the details of the images. Subsequently, the images are stitched using registration parameters obtained by calibration. Finally, the whole system is built on a low-power embedded platform to accomplish fast calibration and mosaic process based on calibration parameters in multi-scene. The experiment results show that the proposed method can accomplish camera calibration quickly and accurately in indoor and tunnel scenarios, and the image mosaic process is time-consuming. Meanwhile, it can ensure better stitching accuracy and imaging effect, and has strong robustness.

Keywords: image mosaic; camera calibration; cylindrical projection; multi-band fusion

Citation: Fu Z Q, Zhang X L, Yu C, *et al.* Cylindrical image mosaic method based on fast camera calibration in multi-scene[J]. *Opto-Electronic Engineering*, 2020, 47(4): 190436

1 引言

图像拼接是将同一场景中具有重叠视场的两幅或者多幅图像组合,以产生一幅无缝全景图或高分辨率图像的过程^[1],经过拼接所获得的图像有较大的视场(field of view, FOV)。绝大部分相机视场角约 $35^{\circ}\times 50^{\circ}$,信息获取受限。因此,通过图像拼接,将同一场景的连续图像序列进行拼接,形成具有更大视场的合成图像,可一次性获取给定视点处的全部视觉信息。该技术在地质勘测^[2-3]、医学影像^[4-5]以及虚拟现实^[6]等多个领域中发挥着重要的作用,其技术优势明显。

国内外研究者对图像拼接进行了大量研究,图像配准是其关键步骤,目前主要有基于频域、基于灰度和基于特征点的三种经典配准方法^[7]。Szeliski^[8]于1996年提出利用图像频域特性进行图像拼接的算法,通过傅里叶变换计算图像间位移的横向功率谱来实现图像匹配,成为全景图拼接的经典算法。Zou等^[9]提出一种基于时空域流形的视频序列图像拼接算法,对视频序列图像拼接十分有效,拼接结果良好。Li等^[10]提出了一种基于尺度不变特征变换(scale-invariant feature transform, SIFT)的算法,在确保拼接全景图像没有缺失的同时,还减少了整个拼接过程的时间。Wu等^[11]在SIFT的基础上针对汽车后视图像系统使用渐变进出的融合算法,计算时间更短、图像细节特征更完整。2016年,Zhou等^[12]针对钢旋转件缺陷检测图像的特征,提出了一种基于坡度概率测度和RANSAC算法,并利用SIFT进行特征跟踪和提取的图像拼接方法,得到无缝、清晰的零件表面图像,为金属零件表面缺陷的自动精确检测奠定基础。Alomran和Chai^[13]通过检测图像重叠区域,自动对齐并融合拼缝以创建无缝全景图像。该算法可对最小5%的重叠面积、水平或垂直旋转以及非固定轴获取的多幅输入图像进行正确拼

接。上述图像配准算法普遍存在计算量大、执行效率低下的问题。对此,有国内研究者提出基于相机标定的方法^[14-16],节省了绝大部分拼接所需的时间,并获得较高的拼接精度。

但目前基于相机标定的拼接方法受场景限制较大,标定过程复杂;且经图像变换后会破坏成像中的共线条件,不利于后续的图像处理和信息归类。为此,提出一种多场景下基于快速相机标定的柱面图像拼接方法,该方法充分利用棋盘格标定板的特征提取精度高的特点,将其置于两两邻接图像的重叠视场中,对采集的图像序列依次进行角点提取、精确化和匹配等预处理,以准确快速求解出待拼接图像间的配准参数,进而迅速完成图像拼接。采用柱面投影和多频段融合方法使图像保持视觉一致性和保留细节信息。系统建立在低功耗嵌入式平台上,实现了可在多场景下相机标定参数的快速获取及准确拼接。经实验验证,该方法标定过程简单快速且准确,拼接精度高,成像效果好,对于在固定先验场景(大范围的视觉实时监控、医疗影像等)及特征点稀少的隧道场景等图像拼接技术上的应用具有积极意义。

2 算法

本算法可分为标定参数获取和柱面图像拼接两个步骤。第一步是利用棋盘格标定板进行快速相机标定,得出相邻图像间的配准参数;第二步先将图像投影到柱面上,再由标定参数进行配准完成图像拼接,获得一幅高质量的拼接图像。

2.1 标定参数获取

标定参数获取分为两个过程(如图1),过程1先对邻接图像进行角点的检测、亚像素精确化及匹配,利用匹配的角点坐标来计算单应性矩阵 H ,再结合焦距评估计算出像素焦距 F 。过程2以该焦距为半径做柱

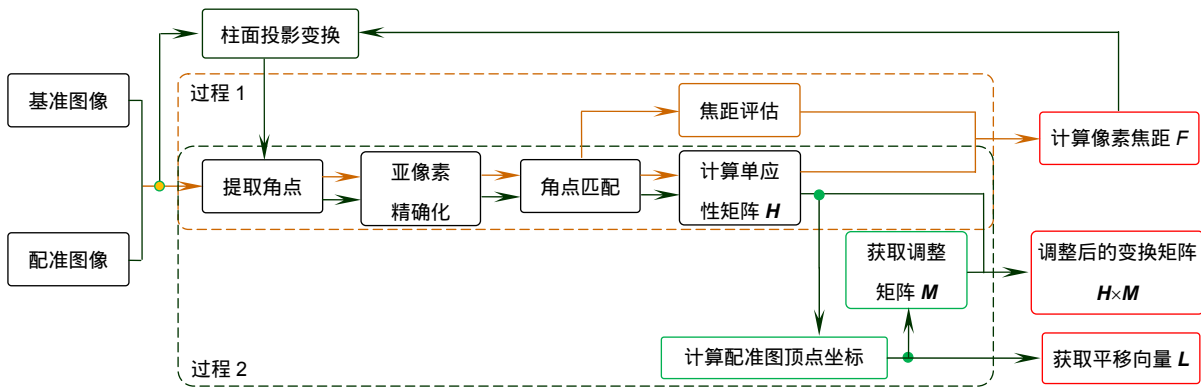


图 1 标定参数获取流程图

Fig. 1 The calibration parameter acquisition flow chart

面投影变换得出柱面图像，对柱面图像再次预处理及计算更新 H ，得出配准图单应性变换后的四个顶点坐标，最后获取调整变换后的单应性矩阵 $H \times M$ 和拼接图像的平移向量 L 。

2.1.1 成像的几何模型

针孔成像模型^[17-18]具有很好的紧凑性，建立几何模型如图 2 所示，被摄体经透镜汇聚于一点(光心)后，其倒影映射到成像平面上。为了便于观察和研究，假设成像平面位于针孔和被摄体之间，被摄体成像方向与实际方向一致。

图中相机坐标系 $X_{cam}Y_{cam}Z_{cam}$ 的原点 O_{cam} 为光心， $X_{cam}Y_{cam}$ 平面与成像平面平行。 Z_{cam} 轴为光轴，光心与光轴的距离即为焦距 f 。图像直角坐标系 $x_{img}y_{img}$ 和像素直角坐标系 uv 都在成像平面上，图像坐标系原点 o_{img} 的像素坐标值为 (u_0, v_0) 。各坐标系间的转换关系为

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \gamma_x f & 0 & u_0 & 0 \\ 0 & \gamma_y f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

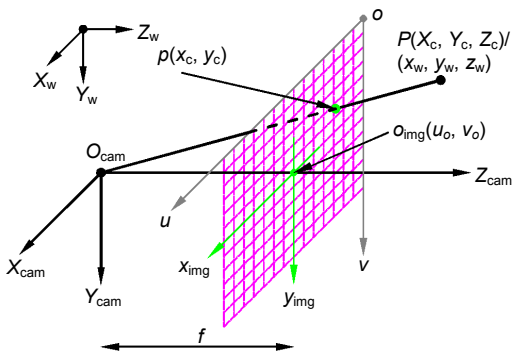


图 2 针孔成像几何模型

Fig. 2 The geometric model of pinhole imaging

$$= KM \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}, \quad (1)$$

式中： $(u, v, 1)^T$ 为点 p 在像素坐标系中的齐次坐标， γ_x 和 γ_y 分别为横轴和纵轴方向上的长度和像素的转换因子， R 为 3×3 旋转矩阵， t 为 3×1 的平移向量， $(x_w, y_w, z_w, 1)^T$ 为三维点 P 在世界坐标系中齐次坐标。 K 为相机内参矩阵， M 为相机外参矩阵。

2.1.2 单应性矩阵计算

根据建立的成像模型，需计算图像平面间像素点间的映射关系，实现图像的配准。平面间的对应关系可通过单应性(Homography)变换^[19]来确定。三维空间中两图像平面间的变换如图 3 所示，相机在两相邻位姿拍摄得到图像 1 和图像 2(配准图像)，可用一个非奇异的 3×3 的矩阵 H 对其映射进行描述，将配准图像平面变换至基准图像平面上。

在图像 1 和图像 2 上分别取齐次坐标点 $A_1(u_1, v_1, 1)^T$ 和点 $A_2(u_2, v_2, 1)^T$ ，点 A_1 和 A_2 的关系可表示为

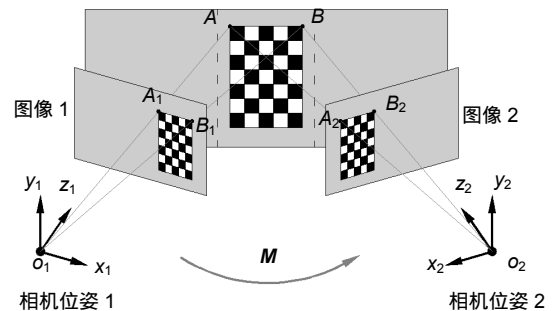


图 3 图像投影变换三维示意图

Fig. 3 3D diagram of image projection transformation

$$c \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = H \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix}, \quad (2)$$

式中： c 为表示尺度因子， H 为单应性矩阵。

单应矩阵 H 只有 8 个自由度，采用线性变换进行求解，将式(2)展开并写成矩阵形式为

$$\begin{bmatrix} -u_2 & -v_2 & -1 & 0 & 0 & 0 & u_1 u_2 & u_1 v_2 & u_1 \\ 0 & 0 & 0 & -u_2 & -v_2 & -1 & v_1 u_2 & v_1 v_2 & v_1 \\ \cdot [h_{11} & h_{12} & h_{13} & h_{21} & h_{22} & h_{23} & h_{31} & h_{32} & h_{33}]^T \\ = A_i h = 0, \end{bmatrix} \quad (3)$$

式中： h 为矩阵 H 的向量形式，因而需求解 8 个未知变量，也就需要 4 对匹配点(每 3 个点不共线)，通常为了精确结果会使用远大于 4 个点来进行计算。此时式(3)中的 A_i 变成一个多行的矩阵 A_n 。因此，需要求解超定方程组 $A_n h = 0$ ，通过齐次线性最小二乘法^[20]对该方程组进行求解，即对 A_n 进行奇异值分解，无需解方程组即可得到 h ，它取值于 A_n 最小奇异值对应的最右侧奇异向量 将 h 重新排列后得出单应性矩阵 H 。

2.1.3 像素焦距计算

为保持拼接图像视觉一致性，在拼接算法中引入柱面投影变换，柱面的半径为相机像素焦距。如图 3，假设 $A(x_w, y_w, z_w)$ 为三维空间点，图像 1 和图像 2 对应像素点 $A_1(u_1, v_1)$ 和点 $A_2(u_2, v_2)$ ，由式(1)可得：

$$\begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = K_1 \begin{bmatrix} R_1 & t_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = K_2 \begin{bmatrix} R_2 & t_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}, \quad (4)$$

式中： $(u_1, v_1, 1)^T$ 和 $(u_2, v_2, 1)^T$ 分别为点 A_1 和点 A_2 的像素齐次坐标， $(x_w, y_w, z_w, 1)^T$ 为点 A 在世界坐标系下的齐次坐标。

相机只做旋转运动，平移向量 $t_1=t_2=0$ 。为了方便计算，假设图像坐标与像素坐标间的平移都为 0，并设 $F_x = \gamma_x f$ 和 $F_y = \gamma_y f$ ，则内参矩阵分别为 T_1 和 T_2 ，外部参数矩阵为 R_1 和 R_2 ，由式(2)可得：

$$\begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = T_1 R_1 R_2^{-1} T_2^{-1} \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix}, \quad (5)$$

式中： $R_1 R_2^{-1}$ 可用相机位姿 1 到位姿 2 的相对旋转矩

阵 R_{12} 代替，该矩阵通过旋转矩阵与欧拉角之间的关系来进行描述^[21]。

由式(2)和式(5)可知：

$$H = T_1 R_{12} T_2^{-1}. \quad (6)$$

进行焦距评估时需要设置图像的长度和宽度相等，即定义两个变量(像素焦距) F_1 和 F_2 ，使得 $F_1 = F_{x1} = F_{y1}$ ， $F_2 = F_{x2} = F_{y2}$ ，由上式变换后可得：

$$\begin{bmatrix} h_{11} & h_{12} & h_{13}/F_2 \\ h_{21} & h_{22} & h_{23}/F_2 \\ h_{31}F_1 & h_{32}F_1 & h_{33}F_1/F_2 \end{bmatrix} = R_{12}. \quad (7)$$

旋转矩阵 R_{12} 属于标准的正交矩阵，它的行向量和列向量都是两两正交的，且每个行向量和列向量都是单位向量，则有：

$$\begin{cases} h_{11}h_{21} + h_{12}h_{22} + h_{13}h_{23}/F_2^2 = 0 \\ h_{11}^2 + h_{12}^2 + h_{13}^2/F_2^2 = h_{21}^2 + h_{22}^2 + h_{23}^2/F_2^2 \end{cases}. \quad (8)$$

上式得到两个 F_2 的值，采用 OpenCV^[22] 图像拼接算法中的焦距评估，比较两个 F_2 分式形式中分母绝对值的大小，如果前者大于后者，则取两个 F_2 的值中的较大者，反之取较小者。

同理可得到相机的位姿 1 处的像素焦距 F_1 。

2.2 柱面图像拼接算法

该图像拼接过程如图 4，是基于标定得到的配准参数。首先采集图像序列，通过设计的自适应工业相机对焦转台装置，根据预设的值依次改变相机的位姿和焦距并触发相机采图，每完成一幅图像采集则进行柱面投影变换，采图和图像变换处理分节点独立进行，效率较高。图像采集数量达到预定值后，采图结束。随后依次读取变换后的图像序列，利用标定得到的变换矩阵 $H \times M$ 进行单应性变换。选取序列中间图像为基准图像，其余的为配准图像，利用标定得到的平移向量 L 拼接图像，再对拼接后的图像利用多频段融合方法消除拼缝线。最后将融合后的图像更新替换为新的基准图像，依次与配准图像拼接融合得到全景图。

2.2.1 柱面投影变换

由于进行图像拼接的原始图像是现实场景在各自观察方向所对应的视平面上的投影，如果对含有重叠区域的原始图像直接进行配准并拼接，使得我们得到的图像序列的二维投影的坐标系产生差异，这将会破坏现实场景的全景视觉属性，得到图像畸变的全景图拼接结果。而柱面投影变换^[23]是将现实场景投影到一个以固定视点为中心的假想圆柱体表面，通过将图像投影到统一坐标系下以像素焦距为半径的柱面上进行

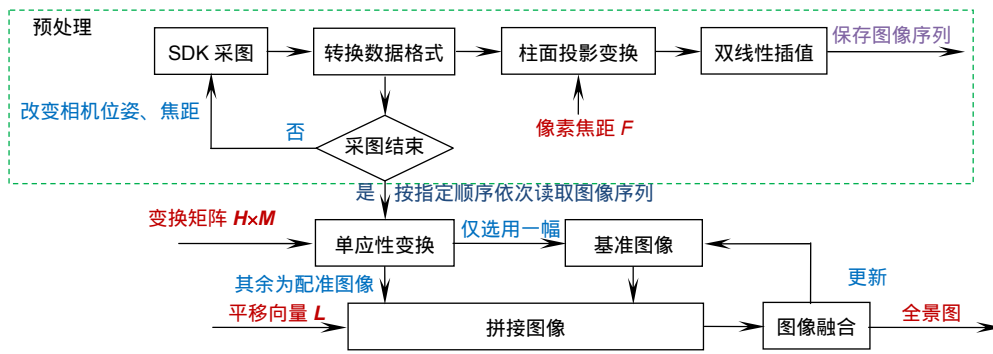


图 4 图像拼接算法框图

Fig. 4 The algorithm framework of image mosaic

拼接，能够维持现实场景中的空间约束关系，从而在后面的拼接过程中保持全景视觉的一致性。柱面投影示意图如图 5 所示。

图 5(a)是三维中成像平面以焦距 f 为半径的柱面投影，图 5(b)为该三维图的俯视图。点 $p(x_c, y_c)$ 为图像坐标系上的一点， x_c 和 y_c 可由式(2)逆变换得出。点 p 经柱面投影变换后为 $p'(x'_c, y'_c)$ 点，它们之间的变换关系可表示为

$$\begin{cases} x'_c = F \times \arctan(x_c / F) \\ y'_c = F \times y_c / \sqrt{x_c^2 + F^2} \end{cases}, \quad (9)$$

式中： F 为相机标定得到的像素焦距， x'_c 和 y'_c 为圆柱面上图像坐标系的横纵坐标值。

而成像处于二维平面上，因此还需将柱面图像反变换至平面上。将点 p' 转换至相机坐标系中：

$$\begin{cases} X'_c = F \times \sin \beta \\ Y'_c = y'_c \\ Z'_c = F \times \cos \beta \end{cases}, \quad (10)$$

式中： X'_c 、 Y'_c 、 Z'_c 分别为至相机坐标系中 p' 的坐标值， $\beta = \arctan(x_c / F)$ ，再根据式(1)和式(2)可以得出

反变换后的像素坐标：

$$\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \lambda' \begin{bmatrix} \gamma_x f & 0 & u_o & 0 \\ 0 & \gamma_y f & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X'_c \\ Y'_c \\ Z'_c \\ 1 \end{bmatrix}, \quad (11)$$

式中： $\lambda' = 1 / Z'_c$ ， $(u', v', 1)^T$ 为点 p' 在成像平面的像素齐次坐标， $(X'_c, Y'_c, Z'_c, 1)^T$ 为点 p' 在相机坐标系中的齐次坐标。

原图像经柱面投影变换以及反变换至平面后的 u' 和 v' 值有大量非整数，而像点坐标要求是整数。采用文献[24]中的双线性插值方法，求解图像变换前后的相应坐标值，完成像素点的匹配，去除图像毛刺，达到好的成像效果。柱面投影变换后的图像如图 6。

2.2.2 单应性变换

单应性变换是将一个平面内的点映射到另一个平面内的二维投影变换，通过标定过程计算图像平面对应关系，将相邻两幅柱面反变换后的图像进行单应性变换，使得相邻图像的重叠部分在二维平面上保持相同尺度。单应性矩阵获取原理及实现如上文标定过

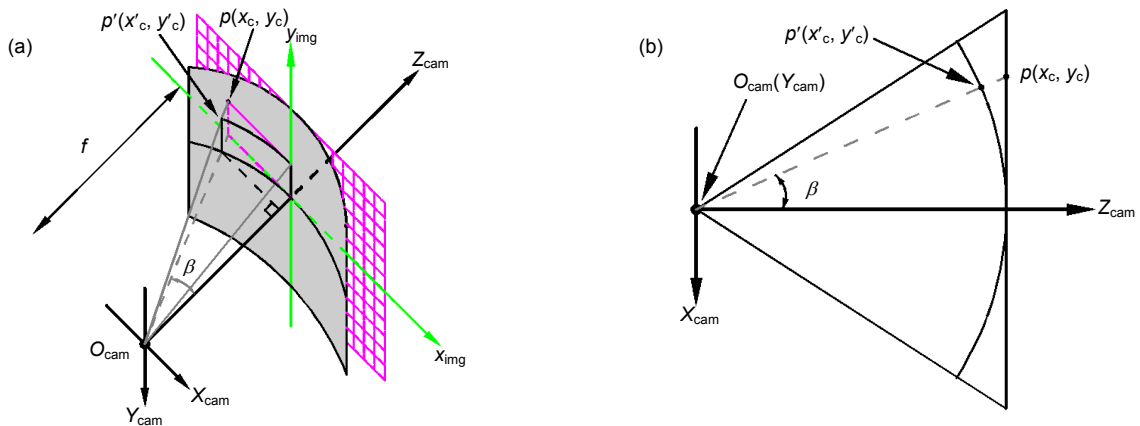


图 5 柱面投影示意图

Fig. 5 The sketch of cylindrical projection transformation

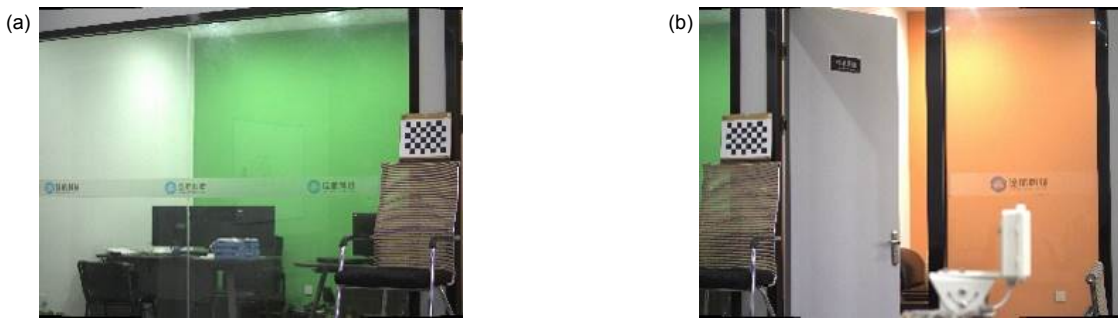


图 6 柱面投影示意图

Fig. 6 The sketch of cylindrical projection transformation

程描述,该矩阵(8个自由度)各参数的作用如表1所示。

2.2.3 拼接图像

相邻图像经透视变换后已处于相同平面,利用标定好的平移向量,将配准图像所有像素点进行平移,与基准图像相应配准像素点重合拼接,未融合的图像拼接如图7所示。图中红框为配准图经透视变换后的边界轮廓,经调整和平移变换后已位于指定拼接位置。

在拼接时选取最右侧图像为基准图像,其左边相邻的图像配准图像,利用标定出的平移向量拼接至基准图上。将拼接好的图像融合后更新为基准图像,重复上述过程可完成拼接。

2.2.4 图像融合

拼接后图像间的衔接不自然,需对其进行融合。像素加权平均法最为简单直接,而在复杂场景容易产生 ghost 区域,可以通过多特征融合的背景建模方法^[25]提高其适应性,但会导致目标局部不完整现象。采用 Brown 和 Lowe^[26]提出的多频段融合方法,通过建立拉普拉斯金字塔(Laplacian pyramid, LP),将待拼接图像分别分解到不同的空间频带上,在各空间频率层上分别进行合并融合,可使各个频带上的特征与细节都保留并融合在一起。

融合过程如图8所示,该融合方法需先建立待拼

表 1 单应性矩阵参数介绍

Table 1 Introduction of homography matrix parameters

参数	作用
h_{11} 、 h_{12} 、 h_{21} 、 h_{22}	表示图像线性变换
h_{13} 、 h_{23}	表示图像平移
h_{31} 、 h_{32}	用于产生图像透视变换

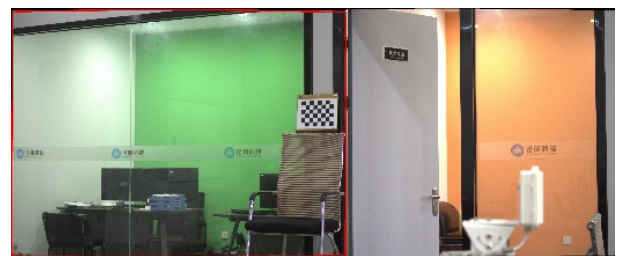


图 7 拼接图像示意图

Fig. 7 The diagram of Image overlay

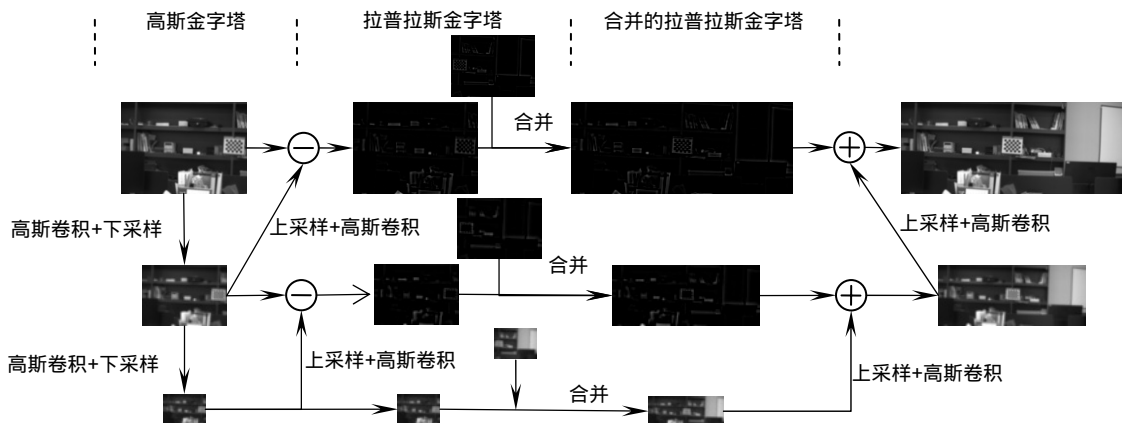


图 8 多频段融合示意图

Fig. 8 The sketch of multi-band fusion

接图像的高斯金字塔,然后用高斯金字塔的每一层图像,减去其上一层的扩展图像(完成上采样并高斯卷积)得到 LP,再将重叠区域的 LP 相同层采用加权平均的方法进行合并。最后将合并后的 LP 从顶层开始进行扩展,将其扩展图像与下一层的合并 LP 图像相加得到下一层的融合图像,逐层递推完成图像重构。

3 实验结果与分析

为在实际环境中验证拼接方法,搭建了如图 9 所示的硬件系统实验平台,分别在室内和隧道场景中进行相机标定和图像拼接,并与其它图像拼接方法进行了拼接时间和拼接精度的对比分析。通过实验对本方法进行误差分析,并功能模块化于隧道巡检机器人中,在杭州文一路隧道和上海四平路隧道试行,进而验证拼接系统的稳定性。

3.1 实验系统的搭建

为了使该方法更具灵活性、通用性,将整个拼接系统模块化,建立在低功耗嵌入式平台上(如图 10)。其既执行舵机运动和相机采图的底层控制,又作为上位机在 Linux 下实现相机标定和图像拼接算法,同时还通过串口和 ROS(robot operating system)实现各功能

模块的通信,从而实现快速的相机标定,且适用于多种场景。

该实验还设计了一套工业相机自适应调焦的装置,镜头调焦圈外接齿轮与舵机转盘上的齿轮啮合,舵机用 L 型连接件固定于相机架上,整个装置随转台转动。这套装置可由嵌入式平台进行自适应调焦,以解决工业相机景深不足的问题。

低功耗嵌入式平台搭载 ARM Cortex-A57 内核;采用大面阵工业相机,分辨率和帧率分别为 4096×3000 和 9 fps,搭配焦距 25 mm 工业定焦镜头;舵机的最小角度 0.088°,转台和调焦舵机的空载转速分别为 63 rpm 和 55 rpm,静止扭矩分别为 6.0 N·m 和 2.5 N·m;结构件采用 3D 打印和钣金加工。

3.2 拼接结果

分别在室内和隧道场景中由该系统拼接完成的全景图如图 11(a)和 11(b)所示,其拼接相关参数如表 2 所示,只需进行一次标定,获取图像拼接过程中所需的各图像间配准参数,进而大量减少图像拼接所需的时间,且该系统建立在低功耗嵌入式平台上,当标定位置 and 姿态与工作点位置和姿态不同时,可快速完成相机标定参数的获取。

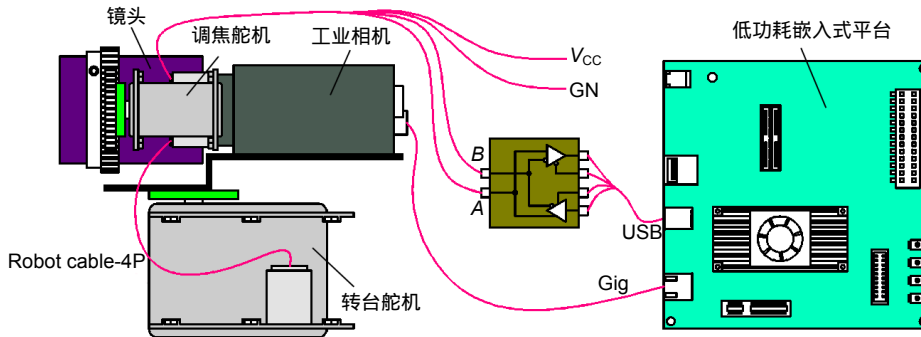


图 9 硬件系统图

Fig. 9 The hardware system diagram

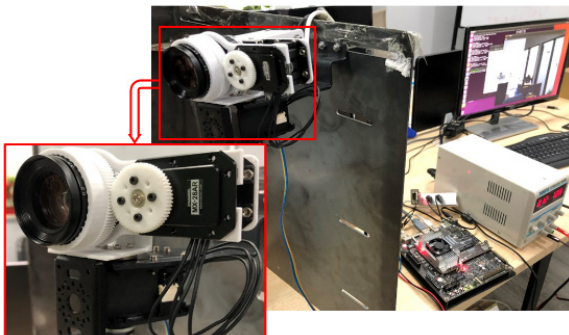


图 10 实验平台

Fig. 10 Experimental platform

表 2 图像拼接过程相关参数

Table 2 The parameters of image mosaic process

名称	参数
单幅图像分辨率	1.23 千万像素
单幅图像视场角(D×H×V)	22°×36°×30°
镜头焦距	25 mm (可调)
相机固定旋转角度	18°
相机相邻位姿间图像重叠度	约 22%
待拼接图像数量	8 幅



图 11 拼接结果。(a) 室内拼接图; (b) 隧道拼接图

Fig. 11 The stitching results. (a) The mosaic image in indoor; (b) The mosaic image in tunnel

3.3 拼接时间对比分析

硬件系统在已完成相机标定的前提下,每采好一张 1200 万像素的图耗时 0.83 s,该过程包括转台转位、舵机调焦和相机 SDK 软触发采图,执行图像拼接算法耗时 1.24 s。从接受拼接指令开始到完成一整套拼接流程总共 9.64 s(以 8 张图为例),各部分所用时间如图 12 所示。整个拼接过程由多个节点独立执行,各部分互不影响。

在相同的硬件设备和图像数据条件下,与文献[4]的利用 OpenCV 和文献[17]基于 SIFT 的拼接方法进行实验对比如表 3 所示,表中数据分别以两幅相邻图像拼接得出,每组数据为 10 次实验平均值。

由于拼接源图像分辨率高,在拼接过程计算量大,从而耗时相对较长。采用低分辨率的 CCD 相机进行实验,通过本文方法完成两幅分辨率为 640 pixels×360 pixels 的图像拼接仅需 0.0113 s,而采用 OpenCV 拼接

方法和基于 SIFT 方法完成拼接分别需要 0.2321 s 和 0.7863 s。综上可知,在相同硬件设备和图像数据条件下,拼接两幅 4096 pixels×3000 pixels 的图像,OpenCV 平均耗时为 1.533 s,基于 SIFT 的拼接方法平均耗时为 8.5488 s,本文方法平均耗时为 0.1854 s,明显快于 OpenCV 和基于 SIFT 的拼接方法。

由于 OpenCV 中的图像拼接算法复杂,对于图像间特征点的检测、匹配筛选和查找最优拼缝线进行融合,都要进行大量的计算,因此需花费大量的时间;基于 SIFT 的图像拼接方法虽然使得图像间配准更精确,但在图像特征检测过程中检测到的特征点数目巨大,图像配准过程耗时长;而本文的方法只需执行相机标定,可快速获取拼接配准参数,随后由标定参数配准图像,极大地提高了图像拼接速度。本文方法的拼接速度是 OpenCV 的 8 倍,而对比基于 SIFT 的拼接方法的运行速度更是有 45 倍,具有非常大的提升。

表 3 图像拼接速度对比

Table 3 Comparisons of image mosaic speed s

	本文方法	OpenCV	基于 SIFT 方法
Image1r_2l	0.1857	1.3945	11.9374
Image2r_3l	0.1773	1.8812	9.1451
Image3r_4l	0.1981	1.3327	6.7098
Image4r_5l	0.1804	1.3343	7.2453
Image5r_6l	0.1851	1.8821	10.1633
Image6r_7l	0.1879	1.5326	7.7876
Image7r_8l	0.1834	1.3769	6.8534

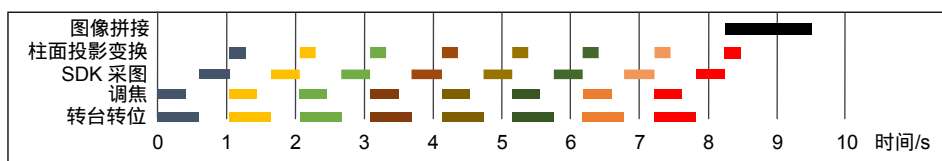


图 12 拼接过程时间柱状图

Fig. 12 The time bar chart of splicing process

3.4 拼接精度对比分析

对含有重叠区域的原始图像直接进行配准并拼接结果如图 13 所示,该结果由 5 幅图像进行拼接融合得到。图像中本来是直线的边缘轮廓,由于直接配准进行变换后产生的非一致性,在拼接得出的全景图中成为折线,破坏了现实场景的全景视觉属性,得到图像畸变的全景图像,成像效果差。

与图 11 对比可知,柱面拼接结果基本无扭曲,能够较好地保持视觉一致性。采用的多频段融合方法使得拼缝自然衔接,无重影和移位等现象,具有较好的成像效果,能够满足实际需求。

3.5 误差分析

本实验的拼接质量依赖于相机标定参数的准确度,而影响标定过程的主要是相机机械安装误差和标定板放置的位置。将以标定后得到像素焦距 f 和拼接

平移值 L_u 作为标定误差指标,它们分别影响图像配准和拼接融合的精度,分别从相机欧拉角中的 pitch(倾斜)和 roll(滚动)旋转方向对其安装位置和标定板在图像重叠视场位置进行误差分析。

本实验通过二维旋转微距云台分别对相机 roll 方向(图 14(a))和 pitch 方向(图 14(b))的误差进行分析,云台底座固定在支撑架上,并在 90° 的位置锁紧使云台旋转活动座保持竖直,以实现相机 pitch 方向下的转动。安装支架与云台快装板连接固定,相机随着云台在 roll 方向下转动,从而达到控制实验变量的目的。

3.5.1 相机 roll 方向下旋转误差

锁紧云台的旋转活动座使相机 pitch 方向保持水平,旋动云台控制相机 roll 方向下的旋转,每旋转 1.5° 取一组数据进行标定,得到对应的像素焦距 f 和平移向量中的横向平移值 L_u , 如图 15 所示。



图 13 无柱面投影变换拼接结果

Fig. 13 Stitching results without cylindrical projection transformation

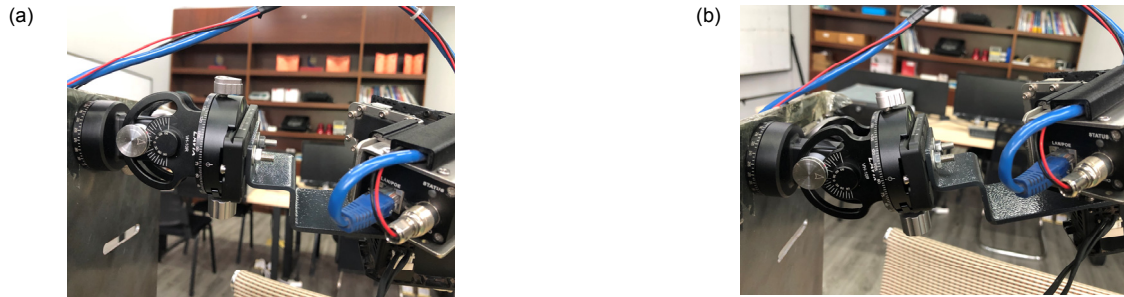


图 14 Roll 方向(a)和 pitch 方向(b)下旋转平台

Fig. 14 Rotary platform in roll direction (a) and pitch direction (b)

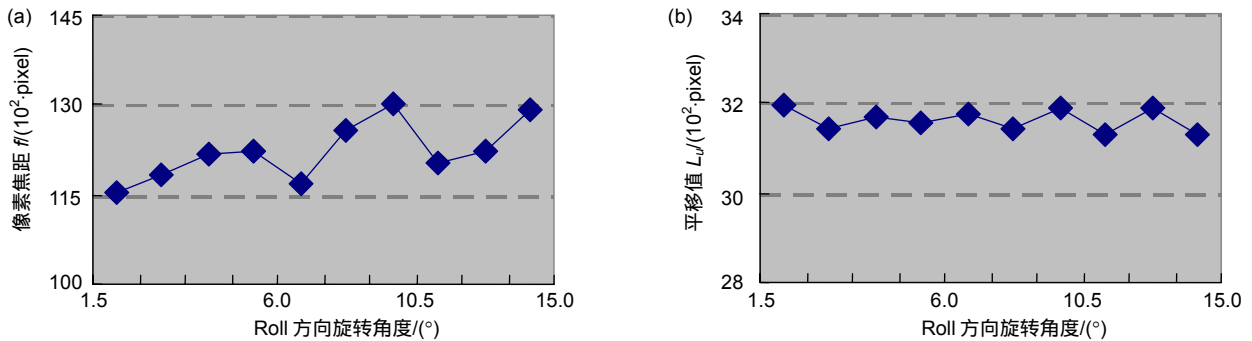


图 15 Roll 方向像素焦距(a)和平移值(b)实验结果

Fig. 15 The experimental results of roll direction pixel focal length (a) and shift value (b)

由图 15 可知, 像素焦距在 11238 pixels 到 13008 pixels 范围内, 最大的误差为 15.75%, 导致在图像配准过程中变换尺度大, 容易造成图 16(a)中的图像畸变, 该图由相机在 roll 方向下旋转 15°标定并拼接得到, 图 16(b)为相机水平实验结果, 图中红框为图像经处理变换后的轮廓。通过对比, 相机在 roll 方向下旋转误差极大影响成像效果, 且多幅图像拼接会进行累加。对于拼接平移值 L_u 保持 3200 pixels 上下微动, 其误差在 3%以内, 即对图像拼接融合影响很小, 能够使图像间自然衔接无拼缝, 保证高的图像拼接精度。

3.5.2 相机 pitch 方向下旋转误差

锁紧云台使相机在 roll 方向保持水平, 转动活动座来控制相机 pitch 方向下的旋转, 每旋转 1.5°取一组数据进行标定, 得到对应的像素焦距 f 和平移向量中的横向平移值 L_u , 如图 17 所示。由图可知, f 的最大误差接近至 9.17%, 而 L_u 的误差保持在 0.8%以内, 表

明相机 pitch 方向下旋转误差对标定过程图像变换有一定影响, 相机在 pitch 方向转过 15°的拼接结果如图 16(c)所示, 该图只有视场角整体发生了变化, 并不影响成像效果, 且对于拼接融合几乎无影响。

3.5.3 棋盘格标定板放置位置影响

单应性矩阵由匹配的棋盘格角点计算得出, 使单应性变换以图像上标定板附近区域为基准进行配准, 而在远离标定板的位置图像变换尺度大, 造成配准误差, 从而导致相同场景像素点不重合的现象。分别改变标定板在两相邻图像重叠视场中横向和纵向位置, 进行标定及拼接实验, 标定得出的数据如表 4, 表中 Δ_u 为棋盘格中心角点位于图像重叠区域的横向像素坐标值, 取值为 640 pixels、500 pixels 和 350 pixels 分别表示标定板位于图像重叠区域的靠右侧、中间和靠左侧, 而 V 值则表示棋盘格在图像重叠区域竖直方向的位置。

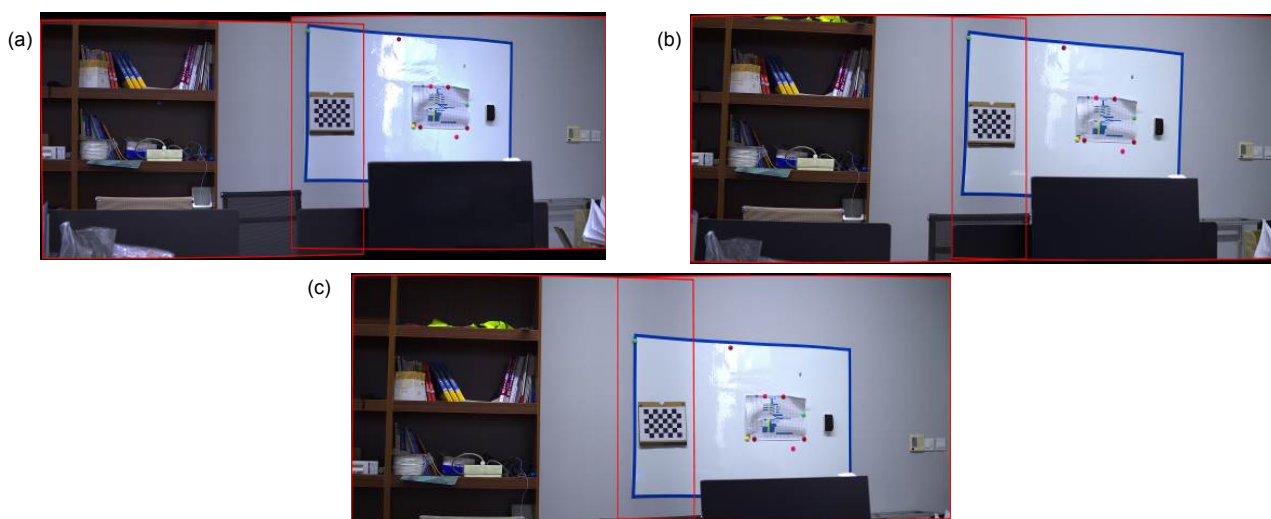


图 16 Roll 和 pitch 方向拼接结果对比

Fig. 16 Comparisons of splicing results in roll and pitch direction

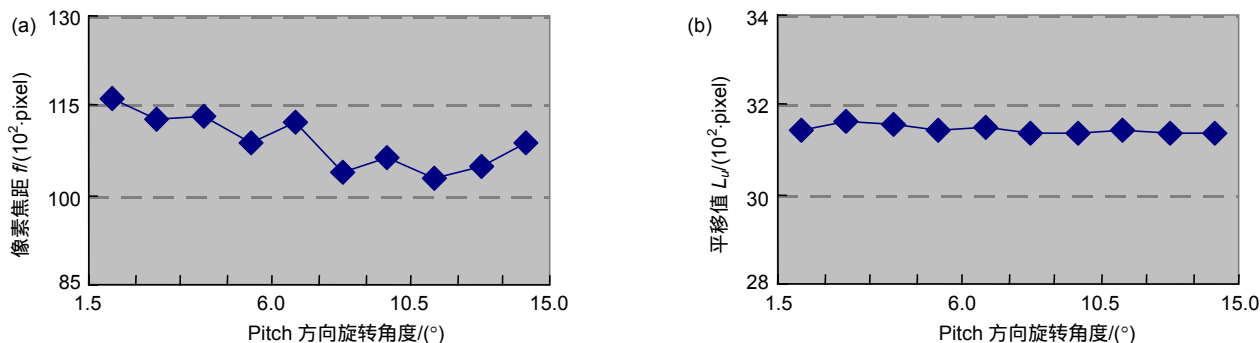


图 17 Pitch 方向像素焦距(a)和平移值(b)实验结果

Fig. 17 The experimental results of pitch direction pixel focal length (a) and shift value (b)

表 4 标定板放置位置误差分析

Table 4 Error analysis of calibration plate placement

	V/pixel	$\Delta_v=640$ pixels		$\Delta_v=500$ pixels		$\Delta_v=350$ pixels	
		f/pixel	L_u /pixel	f/pixel	L_u /pixel	f/pixel	L_u /pixel
1	2600	10859	3147	11809	3145	11757	3147
2	2400	10618	3135	11646	3134	10683	3143
3	2200	12331	3153	11767	3137	11882	3144
4	2000	11288	3145	11423	3145	11567	3140
5	1800	11677	3142	10895	3147	11652	3140
6	1600	11392	3136	10654	3138	11267	3136
7	1400	11523	3145	10739	3148	11516	3138
8	1200	12137	3142	11172	3148	11091	3147
9	1000	12375	3146	10539	3151	11540	3137
10	800	11709	3138	11485	3143	12335	3145
11	600	12193	3139	11592	3141	12506	3144
12	400	11639	3146	11873	3147	12184	3144

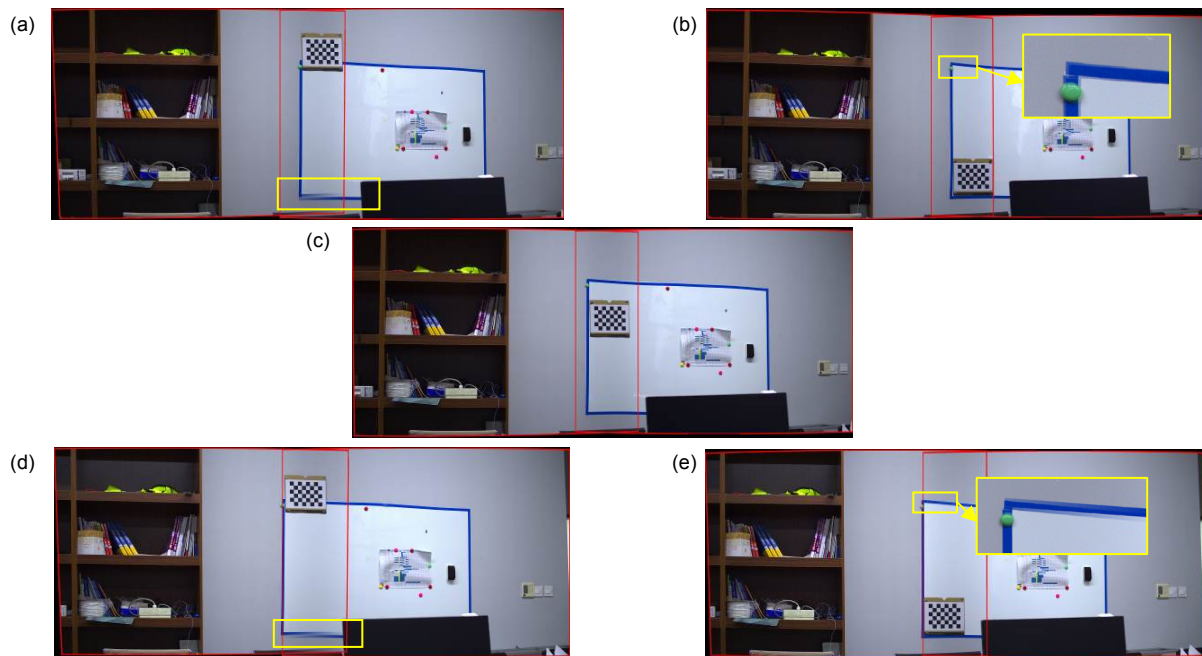


图 18 标定板不同位置的拼接结果

Fig. 18 Stitching results of different positions of the calibration plate

通过表中数据可知，在横向位置 $\Delta_v=500$ pixels 情况下，改变标定板纵向位置使像素焦距值 f 误差最高达到 12.66%，而 Δ_v 取值为 500 pixels 和 350 pixels 时，该误差分别达到 16.55%和 17.06%。即标定板越偏离重叠区域中心图像配准误差越大，使远离标定板的像素点间产生偏移错位(见图 18)。以同样方式分析其对横向平移值 L_u 的影响，误差保持在 0.6%以内，即对图像拼接融合几乎不产生影响。根据实验结果可得出，标定板需位于图像重叠区域中间位置来保证图像配准精度，并实现图像准确拼接与融合，达到较好的成像效果。

4 结 论

基于相机标定的图像拼接方法不受环境光照条件影响，在不失精度的前提下具有较强的鲁棒性，利用标定参数来代替图像配准过程，可有效减少图像拼接所需的时间。本文提出的基于棋盘格的快速相机标定的柱面图像拼接方法在此基础上简化标定过程，标定方便快捷、不受场景限制。采用柱面投影变换，将图像投影在圆柱面上进行拼接，能够保持视觉一致性，具有较好的成像效果。通过室内和隧道(无特征点)场景中实验分析，验证了本文方法的可行性。实验结果

表明该方法能够快速准确获取标定参数,与其它图像拼接方法相比,拼接时间较少,不存在重影和移位现象,具有较高精度。本文方法依赖于标定参数,连接件的加工误差以及安装精度误差会影响标定过程,产生偏差,使得图像变换的尺度过大,影响成像效果。因此,可与多相机结合设计一套稳定的拼接系统,实现对无特征点或环境变化大的实时图像拼接。

参考文献

- [1] Adel E, Elmogy M, Elbakry H. Image stitching based on feature extraction techniques: a survey[J]. *International Journal of Computer Applications*, 2014, **99**(6): 1–8.
- [2] Cheng Y F, Jin S Y, Wang M, et al. Image mosaicking approach for a double-camera system in the GaoFen2 optical remote sensing satellite based on the big virtual camera[J]. *Sensors*, 2017, **17**(6): E1441.
- [3] Bosch J, Gracias N, Ridao P, et al. Omnidirectional underwater camera design and calibration[J]. *Sensors*, 2015, **15**(3): 6033–6065.
- [4] Chen C T, Pan Z W, Shen H L, et al. Image stitching and partitioning algorithms for infrared thermal human-body images[J]. *Opto-Electronic Engineering*, 2019, **46**(9): 180689.
陈晨涛, 潘之玮, 沈会良, 等. 一种人体热红外图像拼接及部位划分方法[J]. *光电工程*, 2019, **46**(9): 180689.
- [5] Kaynig V, Fischer B, Müller E, et al. Fully automatic stitching and distortion correction of transmission electron microscope images[J]. *Journal of Structural Biology*, 2010, **171**(2): 163–173.
- [6] Wu L P, Hu Y. Virtual reality technology of image smoothing in cylindrical panoramic image mosaic[J]. *Science Technology and Engineering*, 2017, **17**(31): 271–276.
吴丽萍, 胡郁. 柱面全景图图像拼接中图像平滑的虚拟现实技术[J]. *科学技术与工程*, 2017, **17**(31): 271–276.
- [7] Seo S, Jeonz S, Lee S. Efficient homography estimation method for panorama[C]//*Proceedings of the 19th Korea-Japan Joint Workshop on Frontiers of Computer Vision*, 2013: 209–212.
- [8] Szeliski R. Video mosaics for virtual environments[J]. *IEEE Computer Graphics and Applications*, 1996, **16**(2): 22–30.
- [9] Zou L H, Chen J, Zhang J, et al. An image mosaicking approach for video sequences based on space-time manifolds[C]//*Proceedings of the 29th Chinese Control Conference*, 2010.
- [10] Li L N, Geng N. Algorithm for sequence image automatic mosaic based on sift feature[C]//*Proceedings of 2010 Wase International Conference on Information Engineering*, 2010.
- [11] Wu Z W, Zhu L R, Sun X C. Expansion of the visual angle of a car rear-view image via an image mosaic algorithm[J]. *Optical Engineering*, 2015, **54**(5): 053101.
- [12] Zhou A W, Shao W, Guo J J. An image mosaic method for defect inspection of steel rotary parts[J]. *Journal of Nondestructive Evaluation*, 2016, **35**(4): 60.
- [13] Alomran M, Chai D. Feature-based panoramic image stitching[C]//*Proceedings of 2016 14th International Conference on Control*, 2016.
- [14] Zhang J L, Sun H X, Jia Q X, et al. Image mosaic algorithm based on camera calibration[J]. *Journal of North University of China (Natural Science Edition)*, 2008, **29**(6): 575–579.
张金玲, 孙汉旭, 贾庆轩, 等. 基于相机标定的图像拼接算法[J]. *中北大学学报(自然科学版)*, 2008, **29**(6): 575–579.
- [15] Wang D, Liu F Y, Chen T E, et al. The method of cylinder panoramic image mosaic based on camera calibration parameters[J]. *Science of Surveying and Mapping*, 2016, **41**(7): 150–154, 143.
王冬, 刘凤英, 陈天恩, 等. 一种相机标定参数的柱面全景影像拼接方法[J]. *测绘科学*, 2016, **41**(7): 150–154, 143.
- [16] Ma J L, Zhang J M, Sun W X. Research on panoramic image mosaic method based on camera calibration[J]. *Journal of System Simulation*, 2017, **29**(5): 1112–1119.
马嘉琳, 张锦明, 孙卫新. 基于相机标定的全景图拼接方法研究[J]. *系统仿真学报*, 2017, **29**(5): 1112–1119.
- [17] Zhang Z Y. Flexible camera calibration by viewing a plane from unknown orientations[C]//*Proceedings of the Seventh IEEE International Conference on Computer Vision*, 1999.
- [18] Hartley R, Zisserman A. *Multiple view geometry in computer vision*[M]. Cambridge university press, 2003
- [19] Chuan Z, Long T D, Feng Z, et al. A planar homography estimation method for camera calibration[C]//*Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation. Computational Intelligence in Robotics and Automation for the New Millennium*, 2003.
- [20] Lenth R V. Least-squares means: the R package lsmmeans[J]. *Journal of Statistical Software*, 2016, **69**(1): 1–33.
- [21] Janota A, Šimák V, Nemeč D, et al. Improving the precision and speed of euler angles computation from low-cost rotation sensor data[J]. *Sensors*, 2015, **15**(3): 7016–7039.
- [22] Pulli K, Baksheev A, Korniyakov K, et al. Real-time computer vision with OpenCV[J]. *Communications of the ACM*, 2012, **55**(6): 61–69.
- [23] Xu Y, Zhou Q H, Gong L W, et al. High-speed simultaneous image distortion correction transformations for a multicamera cylindrical panorama real-time video system using FPGA[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2014, **24**(6): 1061–1069.
- [24] Guo H X, Guo H L, Xie K. Improved bilinear interpolation algorithm based on edge information[J]. *Computer Engineering and Applications*, 2011, **47**(31): 171–174.
郭海霞, 郭海龙, 解凯. 基于边缘信息改进的双线性插值算法[J]. *计算机工程与应用*, 2011, **47**(31): 171–174.
- [25] Guo Z C, Dang J W, Wang Y P, et al. Background modeling method based on multi-feature fusion[J]. *Opto-Electronic Engineering*, 2018, **45**(12): 180206.
郭治成, 党建武, 王阳萍, 等. 基于多特征融合的背景建模方法[J]. *光电工程*, 2018, **45**(12): 180206.
- [26] Brown M, Lowe D G. Automatic panoramic image stitching using invariant features[J]. *International Journal of Computer Vision*, 2007, **74**(1): 59–73.

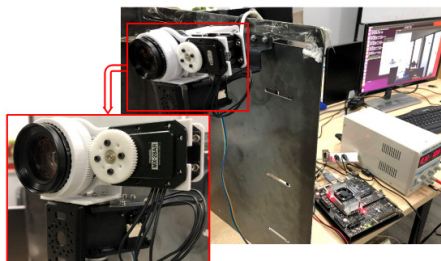
Cylindrical image mosaic method based on fast camera calibration in multi-scene

Fu Ziqiu¹, Zhang Xiaolong², Yu Chen³, Liang Dan¹, Liang Dongtai^{1*}

¹Faculty of Mechanical Engineering and Mechanics, Ningbo University, Ningbo, Zhejiang 315211, China;

²Shanghai Pujiang Bridge and Tunnel Operation Management Co., Ltd, Shanghai 200023, China;

³Ningbo Hermeneutic Aircraft Machinery Technology Co., Ltd, Ningbo, Zhejiang 315200, China



The experiment platform

Overview: Image mosaic is the process of combining two or more images with an overlapping field of view in the same scene to produce a seamless panorama or high-resolution image. The image obtained by mosaic has a larger field of view (FOV). Most of the cameras have a FOV angle of about 35×50 degrees, which limits the acquisition of information. Therefore, through image mosaic, the continuous image sequence of the same scene is stitched to form a composite image with a larger FOV, which can obtain all visual information at a given point of view at one time. This technology plays an important role in many fields, such as geological survey, medical minimally invasive surgery, and virtual reality. Its technical advantages are obvious. Researchers at home and abroad have done a lot of researches on image mosaic, and image registration is the key step. At present, there are mainly three classical registration methods based on frequency domain, gray level, and feature points, respectively. Nevertheless, the above image registration algorithms generally have the problems of large computational load and low execution efficiency. In this regard, some domestic researchers have proposed a camera calibration method, which saves most of the time needed for stitching and achieves high stitching accuracy.

But at present, the mosaic algorithm based on camera calibration is limited by the scene, and the calibration process is complex. The collinear condition in imaging will be destroyed after image transformation, which is not conducive to subsequent image processing and information classification. For this reason, a cylindrical image mosaic method based on fast camera calibration in multi-scene is proposed. This method makes full use of the high accuracy of feature extraction of checkerboard calibration board, which is used to make it in the overlapping field of view of two adjacent images. In order to accurately and quickly solve the registration parameters between the images to be spliced, the image sequence is pre-processed by corner extraction, precision and matching. Then, cylindrical projection and multi-band fusion are used to maintain visual consistency and detailed information. The system is based on a Low-Power Embedded platform, which achieves fast acquisition and accurate mosaic of camera calibration parameters in multi-scene. The experiment results show that the proposed method can accomplish camera calibration quickly and accurately in indoor and tunnel scenarios, and the image mosaic process is time-consuming. Meanwhile, it can ensure better stitching accuracy and imaging effect, and has strong robustness.

The proposed method has positive significance for real-time image stitching without feature points or large environmental changes.

Citation: Fu Z Q, Zhang X L, Yu C, *et al.* Cylindrical image mosaic method based on fast camera calibration in multi-scene[J]. *Opto-Electronic Engineering*, 2020, 47(4): 190436

Supported by National Natural Science Foundation of China (51805280), the Public Technology Application Project of Zhejiang (2017C31094), and Natural Science Foundation of Zhejiang (LQ18E050005)

* E-mail: liangdongtai@nbu.edu.cn