

引用格式: ZHANG Yu, XU Xiping, ZHANG Ning, et al. Research on Visual Odometry Based on Catadioptric Panoramic Camera[J]. Acta Photonica Sinica, 2021, 50(4):0415002

张裕,徐熙平,张宁,等. 基于折反射全景相机的视觉里程计研究[J]. 光子学报, 2021, 50(4):0415002

# 基于折反射全景相机的视觉里程计研究

张裕,徐熙平,张宁,吕耀文,卢毅

(长春理工大学 光电工程学院, 长春 130022)

**摘要:** 为了实现折反射全景相机实时定位,提出了基于特征点法的可以用于折反射全景相机的视觉里程计算法。对折反射全景相机的投影模型重新设计了初始化模块,用于全景相机位姿的初始化。在初始化模块中,提出了适用于折反射全景相机的球面对极几何,通过球面对极几何计算初始化过程中两帧图像间位姿变化关系。为了验证算法鲁棒性和精确度,在相同场景下,将本文算法与使用针孔相机的特征点法的视觉里程计算法进行对比。结果表明,使用折反射全景相机的情况下,室外场景和室内场景下视觉里程计精度分别比针孔相机算法提升了 38% 和 50%。

**关键词:** 视觉里程计;折反射全景相机;球面对极几何;多项式相机模型;本质矩阵

中图分类号: TP242

文献标识码: A

doi: 10.3788/gzxb20215004.0415002

## Research on Visual Odometry Based on Catadioptric Panoramic Camera

ZHANG Yu, XU Xiping, ZHANG Ning, LÜ Yaowen, LU Yi

(College of Optoelectronic Engineering, Changchun University of Science and Technology, Changchun 130022, China)

**Abstract:** In order to realize the real-time positioning of the catadioptric panoramic camera, a visual odometry method based on the feature point method that can be applied to the catadioptric panoramic camera was proposed. the initialization module for the projection model of the catadioptric panoramic camera was designed for the initialization of the camera pose. In the initialization module, a spherical epipolar geometry suitable for catadioptric panoramic cameras was proposed, and the posture change relationship between two frames of images during the initialization process was calculated through the spherical epipolar geometry. In order to verify the robustness and accuracy of the algorithm, in the same scene, the algorithm in this paper was compared with the visual mileage calculation method using the feature point method of the pinhole camera. The results show that, in the case of using catadioptric panoramic camera, the accuracy of visual odometry in outdoor scenes and indoor scenes improves by 38% and 50%.

**Key words:** Visual odometry; Catadioptric panoramic camera; Spherical epipolar geometry; Polynomial camera model; Essential matrix

**OCIS Codes:** 150.1488; 110.2970; 220.3620; 110.4153

## 0 引言

近年来,根据连续的的图像信息来估计相机自身的六自由度运动一直是一个活跃的研究方向<sup>[1-3]</sup>。如今最先进的视觉里程计(Visual Odometry, VO)算法,已经可以精确地、鲁棒地在移动处理器上完成很多有趣

基金项目:国家自然科学基金青年科学基金(No.61803045)

第一作者:张裕(1992—),男,博士研究生,主要研究方向为视觉SLAM。Email: 529957214@qq.com

导师(通讯作者):徐熙平(1969—),男,教授,博士,主要研究方向为光电检测技术与质量控制。Email: xxp@cust.edu.cn

收稿日期: 2020-10-27; 录用日期: 2020-12-28

<http://www.photon.ac.cn>

的应用,比如机器人和无人车等等<sup>[4-6]</sup>,这些都要对自身进行定位。在其他方面还有增强现实(Augmented Reality, AR)和虚拟现实(Virtual Reality, VR)等需要对移动设备进行低延迟和精确姿态估计的应用。

目前已有多种VO算法被提出来,尽管它们的算法流程可能不同,但都有一个共同的前提,就是相邻两帧间必须有足够的重叠区域,只有满足这个条件,单目VO才能正确计算两帧间的相对运动。但是对于常用的针孔相机,由于其视场有限,在相机快速运动或者快速旋转时会导致相邻帧间的重叠减少,这时计算出的位姿不可靠,甚至可能会导致跟踪失败。另一方面,目前的单目VO大多有一个基本假设,即目标场景在处理过程中必须保持静止,因为对动态对象进行的建图是不可靠的,但实际上现实世界并不是一成不变的<sup>[7]</sup>。因此在一些VO算法中使用随机抽样一致算法(Random Sample communication, RANSAC)消除动态对象对算法的影响,然而,当动态对象占到视场的大部分区域时算法也会失效<sup>[8]</sup>。在实际使用视场有限的针孔相机过程中,动态对象占据大部分视场的情况并不奇怪,此时计算出的姿态会不准确,也可能导致算法失效。

由于针孔相机在算法中面对种种问题,本文提出将折反射全景相机应用到单目视觉里程计中。折反射全景相机不需要旋转镜头就可以将360°的场景投影到环形图像上,由于折反射全景相机拥有360°的大视场,在相机移动和旋转很快的时候,仍可以保证相邻帧间有足够的重叠区域;同时,动态物体在单次拍摄中占主导地位的概率极低。因此针孔相机在算法中面对的问题,对于折反射全景相机都可以解决了,并且在现实场景中面对快速运动和动态物体时,算法的鲁棒性也得到了大的提高。

之前也有研究人员做过相关的工作,使用大视场的相机进行视觉里程计和同步定位与建图(Simultaneous Localization and Mapping, SLAM)的研究<sup>[9-10]</sup>。比如MATSUKI H等扩展了DSO<sup>[11]</sup>(Direct Sparse Odometry)的工作,将视场大于180°的鱼眼相机用于DSO算法,证明了大视场相机在用于DSO算法时具有更高的精度和鲁棒性<sup>[12]</sup>。CARUSO D等对LSD算法进行改进,使用未矫正的鱼眼图像序列完成定位与建图工作,并对算法精度和鲁棒性进行了验证,其精度和鲁棒性都有所提高<sup>[13]</sup>。LIN M等使用Ricoh Theta V全景相机设计了VO算法<sup>[14]</sup>。然而Ricoh Theta V相机是由两个鱼眼相机拼接成的360°全景相机系统,在拼接全景图会占用额外的计算资源,在这方面折反射全景相机不用拼接成像显得更有优势。并且相对于大部分鱼眼相机,折反射全景相机的拥有更大的视场,而鱼眼相机想要再增加视场就需要花费更多的成本。因此更有理由将折反射全景相机用于视觉里程计算法中。

本文通过分析折反射全景相机投影模型提出了球面对极几何,用于初始化过程中求解相机运动;另外,针对折反射全景相机模型在ORB-SLAM2的基础上提出了全景相机的初始化算法;并且在数据集上对算法进行了验证,分析了算法的鲁棒性和精确性,并与现有的使用针孔相机的单目视觉里程计算法进行了比较,证明了折反射全景相机的360°大视场的优越性。

## 1 研究方法

### 1.1 折反射全景相机视觉里程计系统

折反射全景相机视觉里程计系统包括折反射全景相机和相应的算法两个部分。折反射全景相机将反射光学器件和折射光学器件结合,通过反射面的反射通常可以采集水平方向360°,垂直方向大于180°视场的全向图。折反射全景成像系统的结构如图1所示,其中包含一个反射镜和一个常规摄像机,反射镜的引入可以大大地提高视场角,并且可以根据应用场合选择合适的面型。常用于折反射全景相机的反射镜面型包括抛物面、双曲面和椭球面等。折反射全景相机具有视场大,结构简单的优势,因此使用折反射全景相机用于定位的视觉里程计算法相较于普通相机的视觉里程计算法也更有优势。

折反射全景相机与针孔相机的结构与投影过程都有很大的差异,所以针对针孔相机的视觉里程计算法在折反射相机上不再适用,要设计适用于折反射相机的视觉里程算法应该从分析折反射全景相机的投影过程与相机模型入手。因此本文针对折反射全景相机的相机模型,在ORB-SLAM2算法<sup>[15]</sup>前端的基础上进行改进,重新设计了ORB-SLAM2的初始化模块,实现了基于特征点法的折反射全景相机的视觉里程计算法,算法流程如图2所示。算法总体分输入模块、预处理模块和位姿解算模块。通过把全景图像序列与标定好的相机内参输入预处理模块对图像进行特征提取与匹配<sup>[16]</sup>,把匹配好的特征点反投影到单位球,然后把反

投影到单位球的特征向量输入位姿解算模块,通过球面对极几何进行初始化和通过PNP算法<sup>[17]</sup>进行跟踪求解,最后输出图像间的旋转矩阵 $R$ 和平移向量 $t$ 。

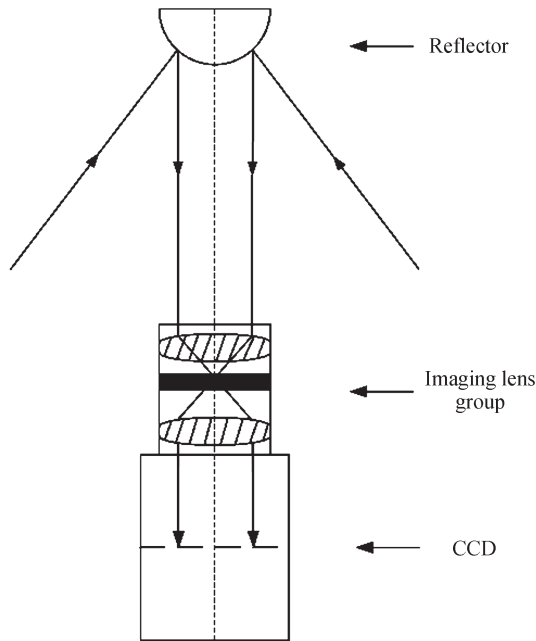


图1 折反射全景相机结构  
Fig. 1 Structure of catadioptric panoramic camera

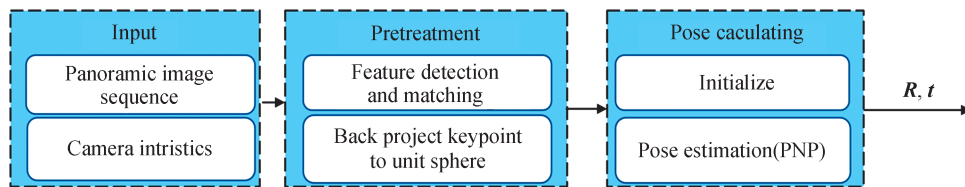


图2 算法流程  
Fig. 2 Algorithm flow

### 1.2 相机模型

使用的是高阶多项式折反射全景相机投影模型<sup>[18]</sup>。该模型广泛应用于折反射全景视觉系统和鱼眼视觉系统中。

图3中上半部分为反射镜,下半部分是投影平面,以视点 $O$ 为原点建立坐标系,其中反射镜是沿 $z$ 轴旋

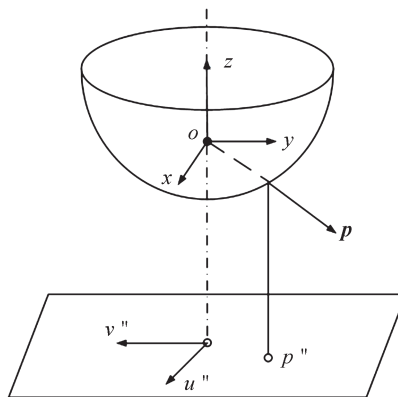


图3 折反射全景相机投影模型  
Fig. 3 Projection model of catadioptric panoramic camera

转对称的。现假设有一空间点  $X$ , 其对应到图像平面的坐标为  $u_i = [u' \ v']^T$ , 对应的像素坐标为  $u_p = [u'' \ v'']^T$ 。而向量  $p$  是视点  $O$  指向空间点  $X$  的向量, 所以空间点  $X$  有对应的空间向量  $p$ 。

从一个像素点到空间点的整个反投影过程如下:

1) 从图像平面到像素平面之间的关系是通过一个仿射变换联系的。可以推出从像素平面上的点到达图像平面的变换关系。

$$\begin{bmatrix} u'' \\ v'' \end{bmatrix} = A \begin{bmatrix} u' \\ v' \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \quad (1)$$

式中,  $A = \begin{bmatrix} c & d \\ e & 1 \end{bmatrix}$ 。

由式(1)可得

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = A^{-1} \begin{bmatrix} u'' - u_0 \\ v'' - v_0 \end{bmatrix} \quad (2)$$

2) 从图像平面到三维空间点的变换关系是通过一个可以表示各种形状的反射镜的广义高阶多项式来联系的。

$$\lambda(u', v', f(\rho)) = \lambda p = X \quad (3)$$

$$f(\rho) = a_0 + a_1\rho + a_2\rho^2 + \dots + a_N\rho^N \quad (4)$$

式中,  $\rho = \sqrt{u'^2 + v'^2}$ 。

根据折反射全景相机的投影模型, 可以使用标定工具箱<sup>[15]</sup>对全景相机进行标定, 从而得到折反射全景相机的内参。其中包括高阶多项式  $f(\rho)$  的系数  $a_0, a_1, a_2, \dots, a_N$  和仿射变换矩阵  $A$  的元素  $c, d, e$ , 以及主点坐标  $u_0, v_0$ 。同时研究数据表明高阶多项式  $f(\rho)$  在四阶的时候其投影效果最好。

### 1.3 全景相机单目初始化

初始化的目的是通过初始两帧图像来计算折反射全景相机的位姿, 并通过三角化计算出一组初始点云地图。单目初始化的大致流程为: 根据视频帧在两帧间进行特征提取和特征匹配, 利用极线约束估计相机初始位姿, 根据相机初始位姿三角化出匹配的特征点的空间位置, 得到世界坐标系下的空间点位置后, 使用 PNP 算法求解运动对特征点进行跟踪匹配, 以此求解相机运动。

在 ORB-SLAM2 中是自动选取符合条件的相邻两帧, 同时估计平面场景下的单应性矩阵  $H$  (Homography Matrix) 和适合普通场景的基础矩阵  $F$  (Fundamental Matrix), 然后根据实际环境自动选择效果好的方法进行单目初始化。

本文针对折反射全景相机的投影模型提出的单目初始化方法是计算本质矩阵  $E$  (Essential Matrix), 利用本质矩阵恢复出初始两帧之间的相对位姿, 并三角化计算出初始的特征点深度, 保存为地图点。然后通过最小化重投影误差的方法联合优化相机的位姿和空间点坐标。本文设计的折反射全景相机单目初始化流程如图 4, 算法具体步骤如下:

1) 选取初始帧。选取特征点数目大于阈值的连续两帧, 进行匹配, 经过 RANSAC 算法<sup>[19]</sup>剔除外点后, 连续两帧的特征点匹配数量仍大于阈值, 选取这两帧进行初始帧。否则重置初始化。

2) 像素点后向投影, 计算空间向量。根据标定得到的后向投影多项式和仿射变换矩阵  $A$ , 把特征点从像素平面反投影到图像, 再转化为单位向量, 用来计算本质矩阵  $E$ 。

3) 计算本质矩阵。在初始帧中随机抽样匹配点对, 并用八点法计算单次采样的本质矩阵, 经过多次 RANSAC 抽样, 计算出最优的本质矩阵  $E$ 。

4) 全景相机位姿求解及三角化。利用计算出的本质矩阵  $E$ , 恢复出折反射全景相机的初始位姿, 并把第一帧图像的相机坐标系作为世界坐标系, 采用直接线性求解 (Direct Linear Transform, DLT) 方法对待估计特征点对应的归一化向量进行三角化, 恢复出对应的三维坐标。

5) BA (Bundle Adjustment) 优化<sup>[20]</sup>。对得到的相机初始位姿和点云地图, 采用光束法平差的方法进行全局 BA 优化, 得到更准确的结果。初始化流程里的 BA 优化是全局的 BA 优化, 对初始两帧的位姿和地图点

进行优化,其实此时虽然是全局优化其实只有两帧而已,是为了得到初始化更加精确的初始位姿和点云地图。

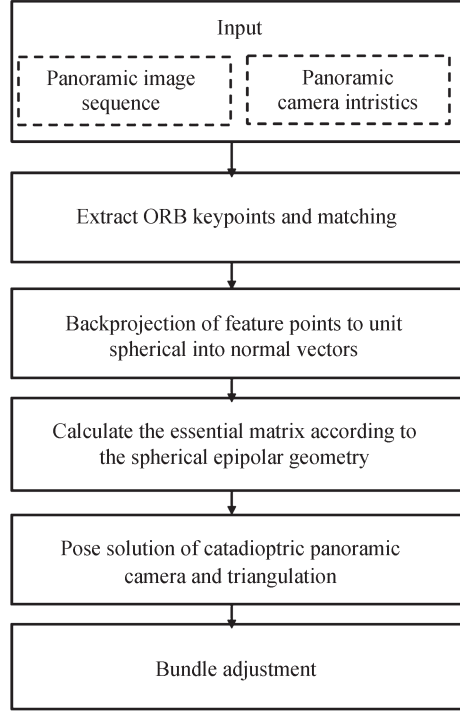


图4 折反射全景相机单目初始化流程

Fig. 4 Flow chart of monocular initialization for catadioptric panoramic camera

#### 1.4 折反射全景相机球面对极几何

以前后时刻两帧图像为基础,提取ORB特征点并进行特征匹配后,对其运动展开预估。1.3节中提出根据本质矩阵 $E$ 来恢复折反射全景相机的初始位姿,在恢复相机位姿的过程中需要用到对极几何的原理,这里结合折反射全景相机的投影模型提出适合折反射全景相机的球面对极几何。

如图5所示,空间点 $X$ 对应到相机 $C_1$ 和 $C_2$ 的球面上的点分别为 $x_1, x_2$ 。向量 $p_1, p_2$ 分别为空间点 $X$ 对应的两个特征点反投影的法向量。

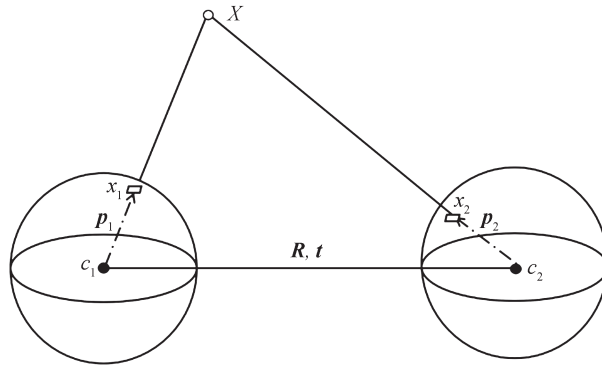


图5 折反射全景相机球面对极几何

Fig. 5 Catadioptric panoramic camera spherical epipolar geometry

为求取相机 $C_1$ 到相机 $C_2$ 的运动,假设他们的运动关系用旋转矩阵 $R$ 和平移向量 $t$ 来表示。设 $X$ 点的空间位置为 $X = [X, Y, Z]^T$ ,两个特征点对应到单位球面的两个法向量 $p_1, p_2$ 与空间点 $X$ 的对应关系为

$$s_1 p_1 = X \quad (5)$$

$$s_2 p_2 = RX + t \quad (6)$$

当采用齐次坐标来表示空间点坐标时,式(5)、(6)具有尺度意义下的相等,可记作 $sp \simeq p$ ,这时有

$$p_1 \simeq X \quad (7)$$

$$p_2 \simeq RX + t \quad (8)$$

经化简得

$$p_2^T t^\wedge p_2 \simeq p_2^T t^\wedge R p_1 \quad (9)$$

式(9)中等式左侧为0,此时就得到了全景相机模型下的球面对极几何公式

$$p_2^T t^\wedge R p_1 = 0 \quad (10)$$

$$t^\wedge = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} \quad (11)$$

在已知 $p_1, p_2$ 的情况下可以求出 $E$ ,然后根据 $E$ 恢复出相机的位姿。

### 1.5 运动估计

经过单目全景相机初始化,得到全景相机的初始位姿和初始的地图点云。接着提取连续帧间的特征点,匹配特征点对并根据帧间模型求解相对位姿变换,本文考虑三种帧间模型,分别如下:

**匀速运动模型:**假设全景相机是匀速运动的,即认为当前两帧间的相机运动与之前两帧间的运动是相同的,可以利用上一时刻的相机位姿和运动速度预测当前时刻的相机位姿。匀速运动模型可以用公式表示为:假设前一帧相机姿态为 $T_1$ ,前一帧到当前帧的变换表示为 $T_{21}$ ,那么当前帧的全景相机姿态初值可表示为 $T_2 = T_{21}T_1$ 。

**参考帧模型:**在运动模型中,若上一帧与当前帧的匹配点数量小于设定阈值,则认为该运动模型失效,此时采用关键帧模型来估计全景相机的位姿。通过特征匹配得到当前帧与参考帧间的二维特征匹配关系,而参考关键帧的大部分特征点在地图中都有对应三维点,可以得到多对匹配的3D-2D点对,计算相机位姿。

**重定位:**全景相机在运动过程中如果出现了跟踪丢失的情况,此时,可以考虑全局重定位继续跟踪相机运动。在关键帧中筛选出与当前帧的匹配点对多余一定数量的作为候选关键帧;利用候选关键帧与当前帧通过PNP算法求解当前帧的全景相机位姿,并选取经RANSAC迭代后匹配内点数量最多的候选关键帧作为匹配关键帧,重新确定全景相机的位姿。

## 2 实验与分析

由于折反射全景相机的特殊结构,导致所拍摄的全景图中只有圆形区域的图像信息是有价值的。在VO算法中,特征点提取与匹配的准确度会直接影响VO的精度,因此在系统中先对全景图像进行预处理就很有必要,对图像中的冗余信息进行掩膜,如图6和7所示,分别对原始图像和对掩膜后的图像进行特征点提取,掩膜后可以剔除大量的边缘错误特征点,从而有利于提升VO算法的精度。

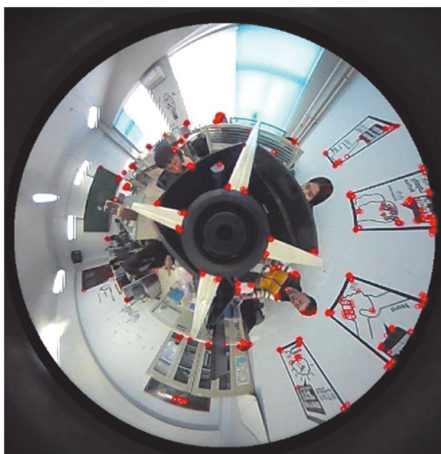


图6 原始图像特征点提取图

Fig. 6 Panoramic image feature point extraction

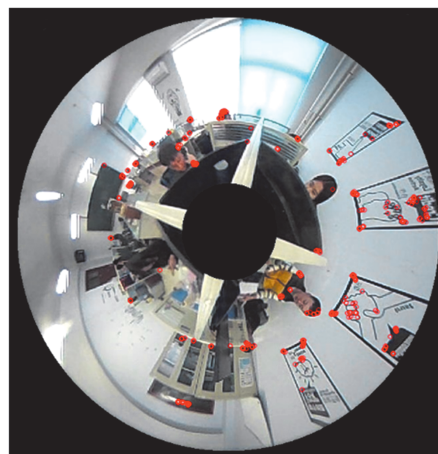


图7 对全景图进行掩膜之后特征点提取图

Fig. 7 Mask the image and extract feature points

为了对所提算法的有效性进行验证,在公开的数据集上进行了实验。本文使用的数据集是 Multi-FoV 数据集<sup>[10]</sup>。这个数据集包含针孔相机、鱼眼相机和折反射相机拍摄的场景数据,可以用来验证折反射全景相机视觉里程计算法的精确度和鲁棒性。该数据集分为室内场景和室外场景两个序列,并且都提供了真实轨迹,用于实验结果的对比和评估。图8是用本文算法在数据集室外场景 Urban canyon 序列和室内场景 Room 序列运行得到的全景相机构建的稀疏点云与关键帧轨迹图。

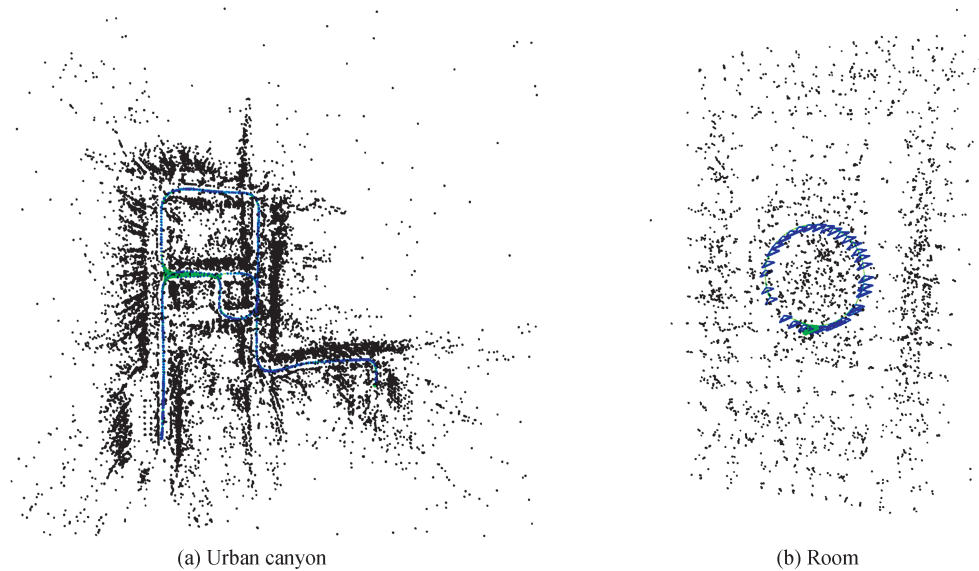


图8 稀疏点云与关键帧轨迹图

Fig. 8 Sparse point cloud and key frame trajectory graph

将本文提出的基于折反射全景相机的视觉里程计算法与当前先进的针孔相机的视觉里程计的精度进行对比。在相同的场景下,分别使用折反射全景相机和针孔相机对运动轨迹进行计算,其中针孔相机使用的算法是现在开源的主流框架的视觉里程计算法,分别对比了直接法的 DSO 算法,和特征点法的 ORB-SLAM2 和 ORB-SLAM3<sup>[21]</sup>。将计算得到的运动轨迹和真实值进行比较,如图9所示。

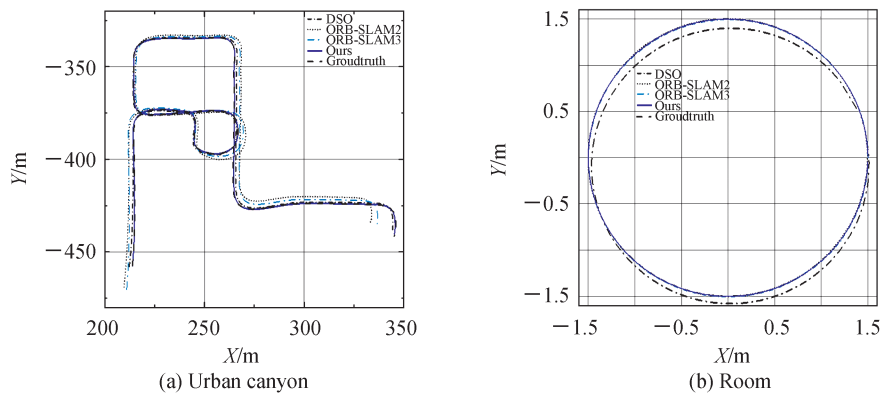


图9 Multi-FoV 数据集各序列下计算轨迹与真实轨迹对比图

Fig. 9 Comparison of calculated trajectory and real trajectory under each sequence of Multi-FoV data set

从图9可以看到本文算法估计的轨迹与真实轨迹基本重合,并且在较小的室内场景中本文算法估计的轨迹与真实轨迹基本完全重合。而 ORB-SLAM2 计算得到的轨迹明显比本文算法得到的轨迹偏移更大。分析图9还可以看出不管是本文算法还是 ORB-SLAM2 算法,在室外场景得到的轨迹都比室内得到的轨迹偏离真值更多,其主要原因是在没有回环检测的情况下随着轨迹总长度的增加,其轨迹误差也会增加;而且对于光照的影响虽然特征点法比直接法更加鲁棒,但光照条件仍然是影响算法精度和鲁棒性的因素之一,室内的光照条件比室外更加均匀也是造成室外轨迹偏差更大的原因之一。

为了对本文所提出的算法精度进行评估,以绝对轨迹误差(Absolute Trajectory Error, ATE)为标准,在相同的场景下分别对本文算法和ORB-SLAM2等主流开源算法计算出的估计位姿和真实位姿间的欧式距离的均方根误差(Root Mean Square Error, RMSE)进行计算,其表达式为

$$\text{RMSE}(s) = \sqrt{\left( \sum_{i=1}^z \|s_i^e - s_i^g\|^2 \right) / z} \quad (12)$$

式中, $s_i^e$ 表示第*i*帧的3D欧氏空间坐标估计值, $s_i^g$ 表示第*i*帧的3D欧氏空间坐标的真实值。并计算了标准方差(Standard Deviation, Std)。记录的各算法运行的RMSE和Std如表1所示。

表1 位姿轨迹误差统计  
Table 1 Pose trajectory error statistics

Sequence	Length/m	Algorithm	RMSE/m	Std/m
Urban canyon	482.59	Ours	0.11	0.04
		ORB-SLAM2	0.18	0.10
		ORB-SLAM3	0.14	0.09
		DSO	0.12	0.07
		Ours	0.002	0.000 7
Room	9.33	ORB-SLAM2	0.004	0.002 0
		ORB-SLAM3	0.004	0.002 0
		DSO	0.020	0.009 0
		Ours	0.002	0.000 7

由表1可知,本文提出的基于折反射全景相机的视觉里程计算法对相机位姿的估计非常精确,在室外场景与室内场景下,绝对轨迹误差与轨迹总长度的百分比分别为0.023%和0.021%。在室外场景下随着轨迹总长度的增加,其绝对轨迹误差要比室内场景下相应的增加,但是绝对轨迹误差与轨迹总长度的百分比仍保持在一个量级。

为了验证使用折反射全景相机的视觉里程计算法在快速运动下的鲁棒性,通过减少数据集序列下的图像帧数来模拟快速运动,在数据集序列中每两帧取出一帧图像作为新的数据进行实验来模拟2倍速度,分别模拟了2倍速度到5倍速度下的跟踪。对各算法正常运行或跟踪失败进行了记录,各算法的跟踪结果如表2。

表2 跟踪结果统计  
Table 2 Tracking result statistics

Algorithm	Speed( $\times 2$ )	Speed( $\times 3$ )	Speed( $\times 4$ )	Speed( $\times 5$ )
Ours	✓	✓	✓	✓
ORB-SLAM2	✓	✓	✓	×
ORB-SLAM3	✓	✓	✓	×
DSO	✓	✓	×	×

表2中对比了本文算法与一些开源算法在不同运动速度下运行的结果,可以看出直接法在4倍运动速度下就已经跟踪失败,ORB-SLAM2和ORB-SLAM3在5倍运动速度下也跟踪失败,而本文使用折反射全景相机的算法一直可以正常运行,证明本文算法拥有更强的鲁棒性。

且本文算法在相同的场景下与使用针孔相机的ORB-SLAM2视觉里程计算法进行对比,在室内和室外场景下本文算法都优于使用针孔相机的ORB-SLAM2视觉里程计算法。本文的视觉里程计算法在室外相对于ORB-SLAM2提升了38%,在室内相对于ORB-SLAM2提升了50%,证明该算法的有效性。

### 3 结论

针对折反射全景相机分析了其投影模型。在该投影模型下提出了球面对极几何,应用于折反射全景相机的初始位姿计算;设计了针对折反射全景相机的单目初始化方法。本文提出的针对折反射全景相机的



VO框架,在室内场景和室外场景的轨迹误差均方根分别为0.002 m和0.11 m,表明本文算法估计的轨迹精确,而且通过分析其绝对轨迹误差与轨迹总长度的百分比,说明了本算法鲁棒性良好。本文算法在相同场景下与针孔相机算法相比时,室外和室内分别精度提升了38%和50%,并且在不同运动速度下进行算法测试,基于折反射全景相机的视觉里程计算法鲁棒性最高。原因在于折反射全景相机对于传统的针孔相机具有更大的视场角,可以获取更多的图像信息,并且在连续帧间的图像重叠更大可以降低跟踪的难度,更不易跟丢,对于提升算法的鲁棒性和精确度方面折反射全景相机更有优势。

本文提出的折反射全景相机的单目VO框架计算出的位姿是没有尺度的,未来应对折反射全景相机的双目VO框架做进一步研究,恢复出位姿变换的尺度。

#### 参考文献

- [1] CHIUSO A, FAVARO P, JIN H, et al. Structure from motion causally integrated over time[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(4): 523-535.
- [2] NISTÉR D, NARODITSKY O, BERGEN J. Visual odometry[C]. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004, 1: 1-8.
- [3] DAVISON A J, REID I D, MOLTON N D, et al. Mono SLAM: real-time single camera SLAM[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(6): 1052-1067.
- [4] SCARAMUZZA D, FRAUNDORFER F. Visual odometry [tutorial] [J]. *IEEE Robotics & Automation Magazine*, 2011, 18(4): 80-92.
- [5] LI S, LI G, WANG L, et al. SLAM integrated mobile mapping system in complex urban environments[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 166: 316-332.
- [6] LI Y, BRASCH N, WANG Y, et al. Structure-SLAM: low-drift monocular SLAM in indoor environments[J]. *IEEE Robotics and Automation Letters*, 2020, 5(4): 6583-6590.
- [7] SAPUTRA M R U, MARKHAM A, TRIGONI N. Visual SLAM and structure from motion in dynamic environments: A survey[J]. *ACM Computing Surveys (CSUR)*, 2018, 51(2): 1-36.
- [8] TAN W, LIU H, DONG Z, et al. Robust monocular SLAM in dynamic environments[C]. *2013 IEEE International Symposium on Mixed and Augmented Reality*, 2013: 209-218.
- [9] JI S, QIN Z, SHAN J, et al. Panoramic SLAM from a multiple fisheye camera rig[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 159: 169-183.
- [10] ZHANG Zhichao, REBECQ H, FORSTER C, et al. Benefit of large field-of-view cameras for visual odometry[C]. *2016 IEEE International Conference on Robotics and Automation*, 2016: 801-808.
- [11] ENGEL J, KOLTUN V, CREMERS D. Direct sparse odometry[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 40(3): 611-625.
- [12] MATSUKI H, STUMBERG LVON, USENKO V, et al. Omnidirectional DSO: Direct sparse odometry with fisheye cameras[J]. *IEEE Robotics and Automation Letters*, 2018, 3(4): 3693-3700.
- [13] CARUSO D, ENGEL J, CREMERS D. Large-scale direct SLAM for omnidirectional cameras[C]. *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2015: 141-148.
- [14] LIN M, CAO Q, ZHANG H. PVO: Panoramic visual odometry[C]. *2018 3rd International Conference on Advanced Robotics and Mechatronics*, 2018: 491-496.
- [15] MUR-ARTAL R, MONTIEL J M M, TARDOS J D. ORB-SLAM: A versatile and accurate monocular SLAM system[J]. *IEEE Transactions on Robotics*, 2017, 31(5): 1147-1163.
- [16] FU Q, YU H, WANG X, et al. Fast ORB-SLAM: a fast ORB-SLAM method with coarse-to-fine descriptor independent keypoint matching[J]. *arXiv preprint arXiv:2008.09870*, 2020.
- [17] LEPETIT V, MORENO-NOGUER F, FUA P. EPnP: An accurate O(n) solution to the PnP problem[J]. *International Journal of Computer Vision*, 2009, 81(2): 155-166.
- [18] SCARAMUZZA D, MARTINELLI A, SIEGWART R. A flexible technique for accurate omnidirectional camera calibration and structure from motion[C]. *Fourth IEEE International Conference on Computer Vision Systems*, 2006: 45-45.
- [19] FISCHLER M A. RANDOM SAMPLE CONSENSUS: a paradigm for model fitting with applications to image analysis and automated cartography[J]. *Readings in Computer Vision*, 1987: 726-740.
- [20] TRIGGS B, MCLAUCHLAN P F, HARTLEY R I, et al. Bundle adjustment—a modern synthesis[C]. *International Workshop on Vision Algorithms*, Springer, 1999: 298-372.
- [21] CAMPOS C, ELVIRA R, RODRÍGUEZ J J G, et al. ORB-SLAM3: An accurate open-source library for visual, visual-inertial and multi-map SLAM[J]. *arXiv preprint arXiv:2007.11898*, 2020.