

# 室内复杂光照环境视觉定位数据集构建研究

刘 春, 陈 晨\*

同济大学测绘与地理信息学院, 上海

收稿日期: 2022年3月21日; 录用日期: 2022年4月17日; 发布日期: 2022年4月24日

## 摘 要

在具有复杂光照特征的室内环境中, 由于缺少公开统一的视觉定位算法验证基准, 机器人的视觉定位任务难以进行验证和评估。因此, 为了打破数据拥有者和使用者之间的鸿沟, 解决面向复杂光照的视觉数据的获取成本高、更新慢的问题, 建立可持续更新的机器人视觉定位数据集是十分重要的。本文聚焦复杂光照的室内环境条件, 探索标准定位数据集组织架构和数据集构建流程, 并根据环境特点进行硬件选型和系统构建, 构建了具有复杂照明特点的视觉定位数据集, 而且基于本数据集评测了多种视觉定位算法。本文为所有的视觉序列提供真实地面位置, 精度达到厘米级, 可以用于测试视觉定位算法处理光照变化的能力。本文提出的构建流程可以为其他场景特点数据集的构建提供可迁移的经验, 本文提出的数据集可服务于视觉定位验证任务, 实现了现实场景和视觉定位算法提升之间的衔接, 能给复杂光照环境下的机器人视觉定位系统提供更多的度量方法。

## 关键词

视觉定位, 复杂光照, 定位精度评估, 机器人

# Data Set Construction in the Complex Illumination Environment for Visual Localization

Chun Liu, Chen Chen\*

College of Surveying and Geo-Informatics, Tongji University, Shanghai

Received: Mar. 21<sup>st</sup>, 2022; accepted: Apr. 17<sup>th</sup>, 2022; published: Apr. 24<sup>th</sup>, 2022

\*通讯作者。

## Abstract

In the indoor environment with complex illumination characteristics, it is difficult to verify and evaluate the visual localization tasks of robots due to the lack of publicly available unified benchmarks for the verification of visual localization algorithms. Therefore, to break the gap between data owners and users and solve the problems of high acquisition cost and slow update of vision data for complex illumination, it is essential to establish a continuously updatable robot vision localization dataset. Focusing on complex illumination environment, this paper explores the standard framework and data set construction process of visual positioning data sets, and conducts hardware selection and system construction according to the characteristics of the environment, and constructs a visual positioning data set with complex lighting characteristics. The dataset evaluates various visual localization algorithms. This paper provides ground-truth locations for all visual sequences with centimeter-level accuracy, which can be used to test the ability of visual localization algorithms to handle lighting changes. The construction process proposed in this paper can provide transferable experience for the construction of other scene characteristic datasets. This dataset serves visual positioning verification task, realizes the connection between the real scene and the improvement of the visual positioning algorithm, and can provide more measurement methods for robot visual positioning system in complex illumination environment.

## Keywords

Visual Positioning, Complex Illumination, Positioning Accuracy Evaluation, Robot

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着人工智能技术的快速深入发展, 智能机器人的研究已经成为一个重要的研究方向。复杂环境下的机器人定位算法是目前智能机器人研究的热点之一。面对层出不穷的算法, 迫切需要创新算法准确性和鲁棒性的验证和评价方法, 而公共数据集是验证定位算法的通用方法。

目前机器人搭载的视觉传感器的价格越来越低, 图像成像的质量越来越高, 在移动机器人的定位传感器选择中优势越来越大。因此, 通过数据集验证视觉传感器的定位精度具有重要的研究价值。近年来, 我国越来越重视新型基础设施建设, 以第五代移动网络为核心的基础设施建设得到了政策的广泛支持。完善数据治理体系, 开发高质量的数据集是我国新基础设施工作的重点之一, 这对数据集相关研究工作具有深刻的指导意义。

在公共数据集上评估算法的准确性和效率是算法获得学术界认可的最直接方式。随着这一思想的普及和算法验证的迫切需要, 机器人视觉定位领域涌现出很多数据集。广泛使用的 KITTI 数据集[1]提供室外场景视觉验证数据, EuRoC MAV 数据集[2]提供室内小场景评价数据。这是两套相对较早、影响较大的算法评测数据。前者基于车辆, 后者基于微型飞行器(Micro Air Vehicle, MAV)。

近年来, 定位算法不断创新, 许多以机器人为载体平台的算法需要在特殊复杂的环境中进行测试, 越来越多的数据集被提出和应用。2009年提出的 New College 数据集[3]提供了以轮式机器人为载体在公园和校园环境收集的视觉定位数据。在密歇根大学北校区收集的 NCLT 数据集[4]的数据涵盖室内和室外

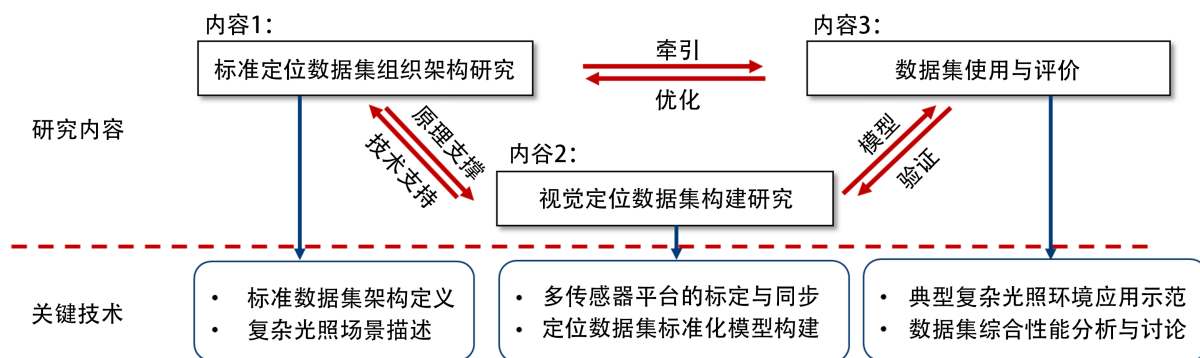
环境。PennCOSYVIO [5]和 TUM VI [6]使用手持设备收集室内和室外场景的高质量序列, 以评估里程表和同步定位与建图(Simultaneous Localization and Mapping, SLAM)算法。ViViD 数据集[7]使用事件摄像机在室内环境中收集数据。Segway DRIVE [8]和 The Newer College [9]数据集考虑了具有挑战性的纹理和照明。然而, Segway DRIVE 数据集只关注了商场场景的序列, 而 The Newer College 数据集采集了开放空间和植被区域的数据。数据集的具体性质如表 1 所示, 目前现有的数据集并没有一个标准框架定义, 而且缺少重点关注弱光照和频繁变化光照环境的数据, 只是在序列中提供几个相似光照条件下记录的少量数据, 不能支持定位算法在光照变化频繁地真实环境中的定位效果验证。

**Table 1.** Comparison of complex illumination visual localization data sets

**表 1.** 复杂光照视觉定位数据集比较表

数据集	年份	数据环境	平台	真实姿势获取方法	包含光照变化	包括序列复杂性度量标准
New College [3]	2009	室外	轮式机器人	九号机器人里程计	×	×
NCLT [4]	2016	室内、室外	轮式机器人	光纤陀螺仪 RTK-GPS EKF-里程计	√	×
PennCOSYVIO [5]	2017	室内、室外	手持	视觉标记板	×	×
TUM VI [6]	2018	室内、室外	手持	起终点的六自由度 动作捕捉系统	×	×
ViViD [7]	2019	室内	手持	动作捕捉系统	√	×
Segway DRIVE [8]	2019	室内	轮式机器人	二维雷达 SLAM	√	×
The Newer College [9]	2020	室外	手持	六自由度 ICP 匹配 定位姿态	×	×

随着视觉定位算法研究的深入, 视觉定位数据集越来越重视所包含的数据能够真实地反应现实世界状况。而传统的定位数据集与定位算法已经难以满足完整、准确的场景理解需求, 因此需要构建适用于复杂光照场景高精度定位的数据集, 实现现实场景和视觉定位算法提升之间的衔接。



**Figure 1.** The research content and key technologies of this paper

**图 1.** 本文研究内容和关键技术

围绕复杂光照环境数据集构建这一关键问题, 本文对定位数据集组织架构、视觉数据集构建流程、数据集的使用和评价三个方面进行了研究, 并构建了具有复杂照明特点的视觉定位数据集, 具体的研究内容和关键技术如图 1 所示。本文对构建机器人视觉定位数据集的过程进行了完整的探索, 为研究人员

后续构建数据集提供可靠的参考。

论文选择同济大学作为实验场地, 在具有复杂照明特点的室内环境中进行视觉定位数据集构建, 如图 2 所示。数据集中记录了具有多样化的场景的轨迹, 例如汽车车灯闪烁和接近完全黑暗的情况, 这写情况经常会在在实际应用中遇到, 而以前的数据集中很少包含这些情况。



**Figure 2.** Data set collection scene: a garage of Tongji University  
**图 2.** 数据集采集场景: 同济大学某车库

## 2. 视觉定位数据集组织架构

视觉定位数据集是涵盖视觉定位算法所需要的场景序列的数据集合。数据集构建者通过搭载视觉定位传感器的平台采集不同特点的场景数据, 并对数据进行加工和同步, 最终形成数据集。数据集选择场景主要依据的原则是高度还原复制现实世界, 因此数据集中通常会选择多个场景或者是在一个场景多次遍历一个位置。在视觉定位数据集中, 这些场景序列的数据主要包含了两种类型: 定位算法输入数据和定位算法验证数据。

### 1) 定位算法输入数据

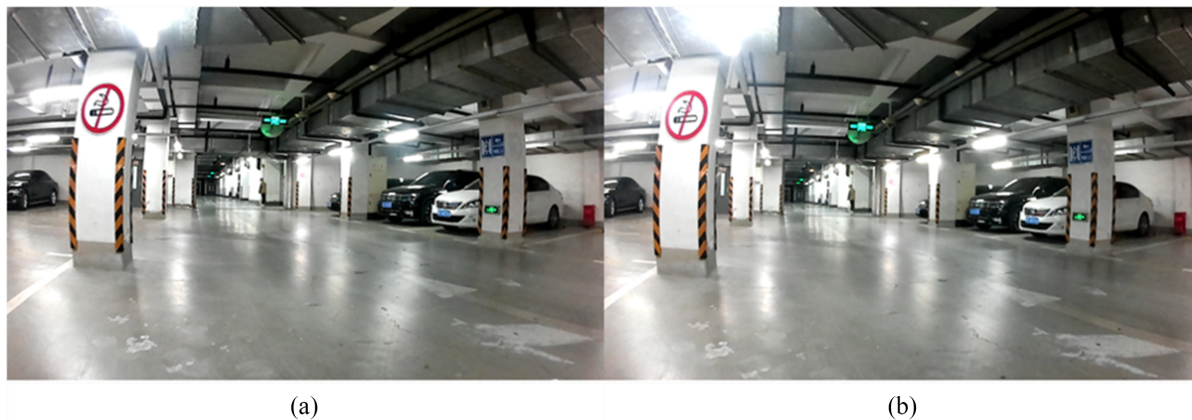
定位算法输入数据是指视觉定位需要的图片数据、视觉传感器的内参数和外参数以及数据集构建者添加到其中的特征, 如物体语义特征等。

数据集中定位算法输入数据的获取涉及到数据集搭建研究人员进行的硬件配置, 包括移动平台和感知传感器的选择。移动平台的主要选择有手持载体、移动背包、轮式机器人及小型无人机。不同传感器配置在实践中的表现仍需探索。因此移动平台可以根据不同环境需求配置不同的视觉感知传感器, 如单目相机、双目相机、事件相机及深度相机, 甚至可以配置辅助视觉定位算法的其他传感器, 如 IMU 等。如图 3 所示, 是双目相机在同一时间的输入数据。

一般情况, 定位算法的输入数据会以图片形式(包括.png、.jpeg、.jpg 等常见格式)或者 ROSBAG 形式提供。每一帧图像数据都会附上拍摄的时间戳信息。

由于视觉定位算法在面对不同环境时运行效果不同, 由此可以给不同采集条件的场景序列进行环境复杂度定性研究, 以此反映不同场景序列能够稳定运行定位算法的困难程度。

传感器间的位置关系通过外参数给出。传感器的镜头模型、内参数、外参数等信息都会由数据集制作人员提前标定并给出。

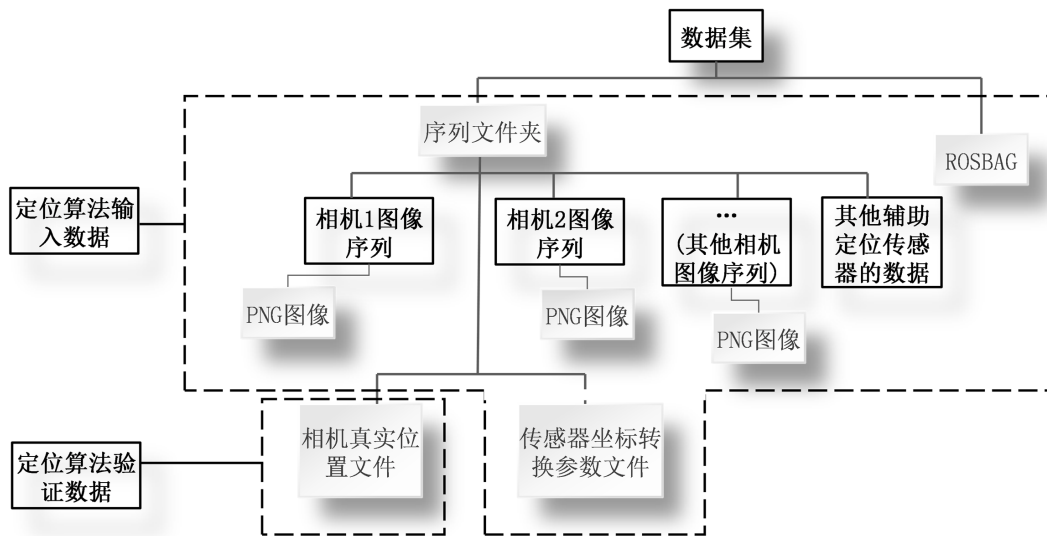


**Figure 3.** Image data collected by the binocular camera at the same time, (a) is the data collected by the left camera, (b) is the data collected by the right camera

**图 3.** 双目相机在同一时刻采集的图像数据：(a) 左相机采集数据，(b) 右相机采集数据

2) 定位算法验证数据

定位算法验证数据是用作与定位算法结果进行比较的较为准确的结果，主要包括位置和姿态两个维度。根据获取数据集的场景不同，获取位置和姿态的技术也有所不同。进行姿态测量的技术主要包括动态捕捉系统、多传感器融合算法、具有 RTK 校正的双天线 GNSS 等技术。在室内环境中，尽管动态捕捉系统工作效率高、精度高，但是面向大面积的室内环境，这个解决方案布设成本太高，无法被数据集制作者广泛使用。在室内环境中，目前被广泛使用的验证数据获取设备包括追踪仪和激光雷达等。



**Figure 4.** Visual positioning dataset organizational architecture

**图 4.** 视觉定位数据集组织架构

视觉定位算法输入数据和视觉定位算法验证数据通过一定的组织形式共同构成完整的视觉定位数据集，具体的视觉定位数据组织架构示例如图 4 所示。尽管不同数据集针对不同的算法验证需求，选取的

场景序列各异, 而且提供的组织形式或者数据格式会有所区别, 但是所有的视觉定位数据集都会具备这两种数据。

正如引言中所说, 现在的视觉定位面临从明到暗或者光照频繁变化的现实场景挑战, 这些场景很容易使现有的视觉定位算法失灵, 将这些场景定义为复杂光照场景。面临缺少针对室内复杂光照的环境数据集的数据困境这一问题, 本文根据数据采集需求搭建数据采集平台, 从而构建了具有复杂光照条件的视觉定位数据集, 提供视觉定位算法的实验和验证数据, 并且基于此数据集评测了多种视觉定位算法。

### 3. 数据集构建流程与关键技术

本文提出的数据集构建方法, 考虑采集的复杂环境特点, 从选择传感器到搭建硬件平台, 从构建软件计算传感器转换参数到提供真实位置参考值, 最终构建出能供算法进行测评的室内复杂环境数据集。本章主要从硬件选型与集成、参数获取与真值设计、数据标准化处理三个方面介绍数据集构建流程。

#### 3.1. 硬件选型与集成

数据集构建工作首先需要完成硬件系统的搭建与外部多源数据的采集和存储。本文构建的数据集基于自主搭建的机器人平台, 根据室内环境多光照变化、通道窄长等特点定制搭载短基线视觉感知传感器, 并对硬件进行模块化管理, 存储采集的图像, 以供定位算法进行输入, 进而实现定位。

数据集构建硬件系统的主要包含以下设备: 视觉传感器、激光传感器、照明系统、定位系统和控制平台。所有设备都连接到一个由碳板和螺钉制成的轮式机器人上, 配备电源和机载计算机用于记录和存储数据, 具体数据采集平台如图 5 所示。



Figure 5. The hardware system of the dataset  
图 5. 数据集的硬件系统

数据集中包括的传感器有以下几种(具体特性见表 2)。

### 3.1.1. 视觉传感器

考虑到室内有部分环境呈现狭长、空间较为拥挤的特点, 选取的传感器要保证光学系统成像清晰, 做到透光性强、杂散光少, 适用性强同时满足高精度和高速度获取图像的要求。因此, 视觉数据考虑使用市面上消费级的多个相机组合获得, 本数据集选择了广泛使用的消费级传感器 MYNTEYE (小觅) 相机 S2110。MYNTEYE 相机 S2110 是一个室内外通用, 精度高, 黑暗环境适用的相机系统, 它搭载了 2 个 200 万像素( $2560 \times 800$ 、 $1280 \times 400$ )的鱼眼相机及 1 个消费级 IMU, 分别能够以 30 Hz、200 Hz 频率捕获视觉、IMU 数据。

### 3.1.2. 激光雷达

真值获取设备是整个数据集采集系统可行性的基础。考虑室内环境的纹理条件相对较为丰富, 使用激光雷达可以获取质量较高的数据, 因此针对室内环境的光照变化较多的条件, 使用对光照不敏感的激光雷达作为获取真值的设备。激光雷达与相机接入同一台计算机内, 实现在计算机内通过系统发出指令同时实现图像采集工作和定位工作, 实现了采集系统的自动化开启。本数据集选用 Hokuyo (北洋) UTM-30LX 激光雷达, 视距长达 30 米, 视场角打到  $270^\circ$ 。为了防止信号遮挡和方便传感器空间信息标定, UTM-30LX 激光雷达水平安装在机器人平台的正上方。

### 3.1.3. 照明设备

由于复杂环境中, 有较多光照亮度变化的情况, 导致在采集视觉定位算法所需要的视觉数据时通过视觉传感器所采集的光线不充分, 所以在数据集采集平台中加上可以提高光线通过量用的设备, 以保证获取的视觉定位算法输入视觉数据的质量, 构造不同照明条件的序列。综合考虑价格、光源价格、损耗和获取照明稳定性等因素, 考虑在机器人平台上组合使用 2 个白光 LED 灯, 与视觉传感器的镜头方向平行进行调节光源。

**Table 2.** List of sensors and their parameters

**表 2.** 传感器列表及其参数

传感器类型	参数
MYNTEYE 相机	鱼眼相机: RGB $1280 \times 400/2560 \times 800@10/20/30$ fps; 像素尺寸: $5.86 \mu\text{m}$ ; 视场角: 垂直方向 $50^\circ$ , 水平方向 $95^\circ$ , 对角线方向 $112^\circ$ IMU @200Hz
Hokuyo UTM-30LX	视距: 30 m; 视场角: 水平方向 $270^\circ$
照明光源	2 个 LED 光源

## 3.2. 数据集构建关键技术——参数获取与真值设计

传感器一体化集成平台的构建是数据集构建的基础。由于该平台集成了不同的传感器, 传感器的安置位置与姿态不同, 获取数据的空间基准也有所不同, 因此在数据集构建过程中要完成空间同步与时间同步。

### 3.2.1. 空间同步——多传感器标定

空间同步是指完成传感器之间的两两标定, 确定传感器间的刚性空间关系。该数据集中涉及多种类型传感器的多模态相对姿态校准, 包括左相机到右相机, 鱼眼相机到 IMU, 鱼眼相机到激光雷达。为了

校准内外参数, 在数据集构建时, 制作了一个 kalibrtag [10] [11]  $30 \times 30$  cm 的浮法玻璃基板, 如图 6 所示, 用传感器采集的标定板图案记录 ROSBAG, 使用 Kalibr 工具箱相关代码[12] [13]。在制作 ROSBAG 的过程中, 将标定板从不同角度旋转, 以达到更好的标定效果[14] [15]。数据集中提供焦距、畸变参数等相机校准结果, 并提供两个相机的相对姿态。由于激光雷达测量的是自身的位姿, 而不是相机的位姿, 因此使用 Cam Laser Calibra Tool [16]完成视觉传感器和激光雷达的平面匹配, 得到了激光雷达到相机的外部标定。

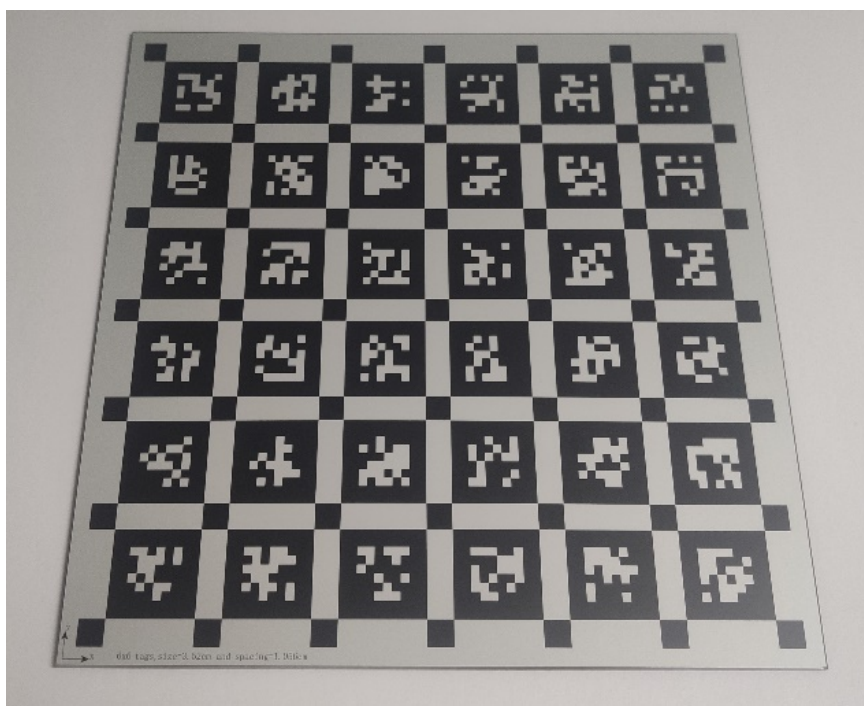


Figure 6. Kalibrtag float glass calibration plat  
图 6. Kalibrtag 浮法玻璃标定板

### 3.2.2. 时间同步

时间同步是指多个传感器之间的时间戳对齐, 具体的时间同步图如图 7 所示。2 个鱼眼相机和 IMU 通过内置芯片进行硬件同步, 误差保持在 5ms 以内。激光雷达和视觉传感器采集数据的方式不同, 激光雷达数据用于提供地面真实位姿数据, 因此假设每个序列激光雷达数据有固定的时间偏移量, 采用最小化绝对轨迹误差(Absolute Trajectory Error, ATE)的均方根误差(Root Mean Squard Error, RMSE)的方法确定不同传感器之间的最佳时间偏移量, 以达到同步轨迹的目的。相机和 IMU 的轨迹相对激光雷达计算得出的轨迹来进行绝对轨迹误差 ATE 的同步。

为了降低定位算法计算时产生的累计误差和测量过程中产生的噪声, 数据集构建和处理数据时, 在每个数据序列采集的起始阶段, 都对用于同步处理的部分数据进行设计。数据集采集过程的初始阶段, 选取纹理较丰富、光照条件良好的环境, 通过对机器人平台进行控制低速运行, 获得稳定鲁棒的数据, 用起始阶段的数据来进行同步处理。由于最小化绝对轨迹误差 ATE 的方法本身受到尺度的影响较大, 如果两条数据输入轨迹的尺度稍有不同, 就会引进系统性误差, 因此在本数据集中不使用整条轨迹数据进行同步处理, 而是只选择起始阶段的低速运行的数据来进行时间偏移量的确定, 避免造成时间偏移量的不准确。通过确定最佳时间偏移量的估计一致性来对同步数据的质量进行评估。



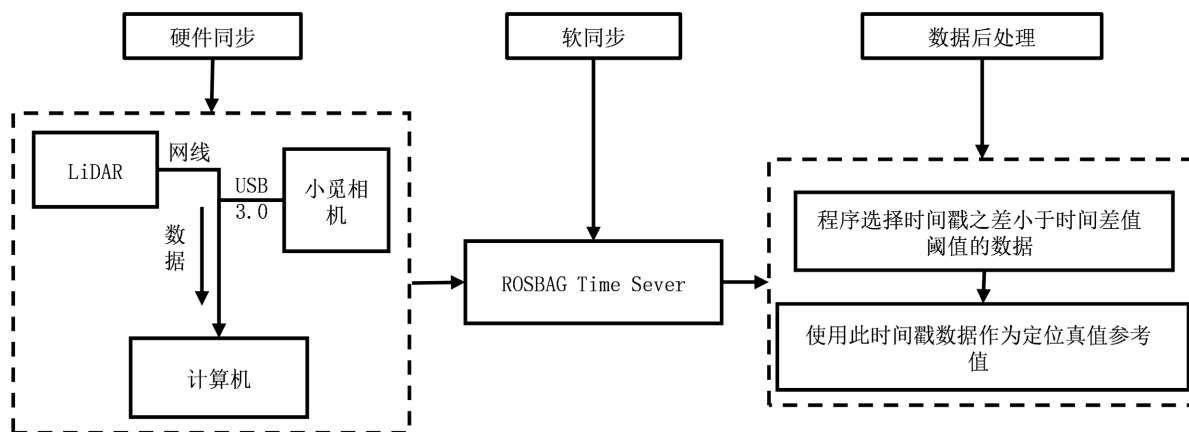


Figure 7. Flow chart of sensors time synchronization  
图 7. 传感器时间同步流程图

### 3.2.3. 数据集真实位姿值获取

由于数据集构建的实验场景场景面积超过 1000 平方米, 本文没有像文献[2]中使用动作捕捉系统来生成实际真实的位姿数据, 而是使用精度较高的 LiDAR SLAM 方法交叉校正 LiDAR SLAM 作为参考位姿, 具体实例见图 8。数据集是在平坦地形中收集的, 因此可以高度变化可以忽略不计。通过激光雷达获取单线激光数据, 并通过文献[17]中提出的交叉校正 LiDAR SLAM 方法进行高精度位姿计算, 以获得每个相机的相对位姿精度。该数据集中的序列以 30Hz 的频率提供经过 2D-SLAM 算法处理后的位姿数据, 定位精度达到了厘米级。通过单线激光雷达数据获取模块能够提供相对高频率高精度的位姿估计。

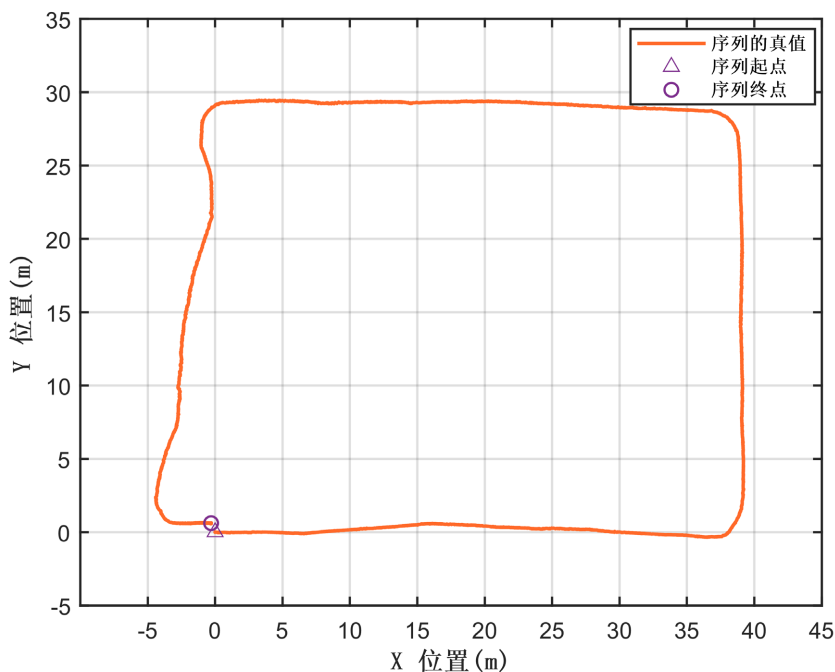


Figure 8. The results obtained by cross correction of LIDAR-SLAM method in Indoor\_HARD\_002 sequence

图 8. 在 Indoor\_HARD\_002 序列中通过交叉校正 LiDAR-SLAM 方法计算得出的结果

### 3.3. 数据标准化处理

#### 3.3.1. 数据采集环境与环境复杂度特征

本文选择同济大学地下停车场一层作为记录站点, 实验环境如图 9 所示。试验场占地面积约 7000 平方米, 环境空旷, 标志物较少, 光线条件较为复杂, 包括人工和自然日光的混合光照条件。此环境可进行视觉定位算法的压力测试。



**Figure 9.** Representative image of the underground parking  
**图 9.** 地下停车场的代表图像

根据定位算法在实际应用中所面临的挑战, 并结合影响定位精度的光照条件对环境进行分级。“复杂光照”只是对环境条件的总结, 不同的光照水平对算法的影响是不同的。在该数据集中, 将复杂照明分为外部照明效果和内部照明效果。外部光线分为强光、中光、弱光和频繁变化四个层次。内置光源是指机器人设备有自己的光源, 包括机器人提供光源和不提供光源两种情况。根据环境的复杂光照条件分级, 该数据集包括 4 个序列, 具体序列属性见表 3。

**Table 3.** Properties of data set sequences

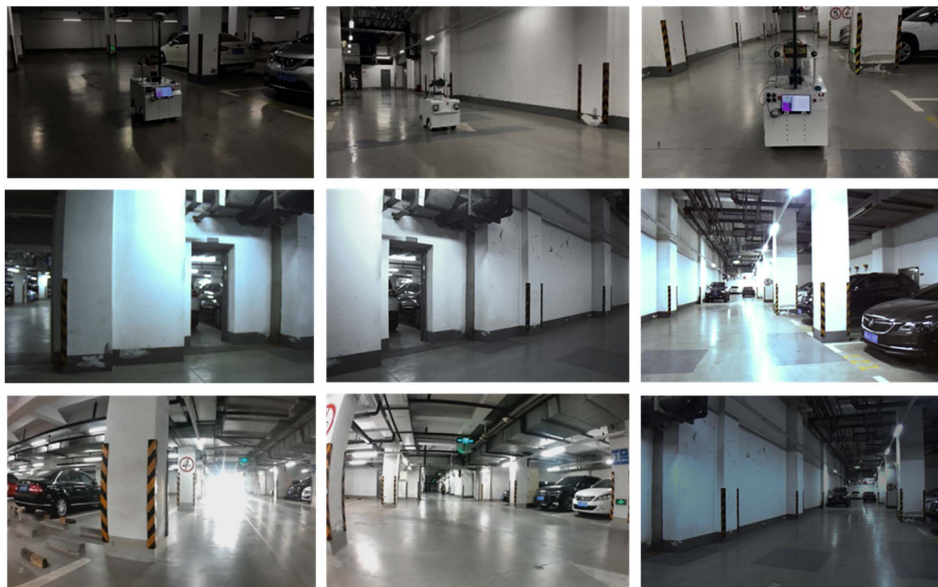
**表 3.** 数据集序列的性质

序列	外部光照	自身光照	平台行驶速率	时间	真实位姿	图像数量 (张)
Indoor_EASY_001	强光	OFF	快速	312s	提供	26,134
Indoor_EASY_002	中等光源	ON	慢速	423s	提供	30,716
Indoor_HARD_001	弱光	OFF	中等	300s	提供	25,848
Indoor_HARD_002	频繁变化	ON	中等	240s	提供	19,826

该数据集提供不同光照条件环境的图像数据, 数据示例如图 10。

#### 3.3.2. 序列描述与文件结构

在数据集构建流程中, 数据集构建者通过固定框架形式组织每个序列里数据, 以方便视觉定位算法研究人员进行算法验证。数据集中的每个序列包含获得传感器数据的集合, 包括定位计算所需的数据和真实位姿数据。参考数据集的标准框架, 使用类似于 EuRoC 数据集格式组织并构建数据集。



**Figure 10.** First line shows a positioning robot doing tasks in an underground parking lot; Second line shows different structural characteristics; Third row shows the obvious lighting changes made  
**图 10.** 第一行显示一个定位机器人在地下停车场的场景；第二行展示了地下停车库不同的结构性特点；第三行显示数据集中的照明变化场景

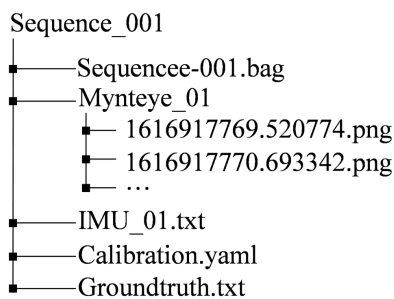
#### 1) 定位算法输入数据

定位算法输入的数据包括标定参数文件、图像文件、IMU 数据。标定参数文件描述传感器相关的所有属性，所有的参数文件中都包含传感器的名称以及一个描述传感器之间位置关系的  $4 \times 4$  齐次变换矩阵。参数文件以 YAML 文件格式给出。图片数据以两种数据格式提供给用户，一种是 ROSBAG 格式，一种是图片格式(.png)。每个相机都有一个单独的文件夹，文件以精确到小数点后六位的时间戳命名。

#### 2) 定位算法验证数据

定位算法验证数据是指数据集能够提供的真实位姿文件，它以 TXT 文件格式给出，文件中包含处理过的所有时间的机器人运行轨迹，文件由多行组成，每行信息都包括时间戳和相应的空间坐标(欧几里得空间系统)和姿态(四元数)。

具体的数据集格式如图 11 所示。



**Figure 11.** A structure diagram of the data set

**图 11.** 数据集结构图

## 4. 数据集使用与评价

为了评价本文发布的数据集，本文使用激光雷达提供的位姿数据作为真实位姿并结合视觉定位算法

的运行结果进行评估。参考 TUM 数据集的方法, 用户只需要提供一个 TXT 文件, 表示通过算法计算出给定序列的轨迹, 每一行包含时间戳及三自由度的位置或者六自由度的姿势。本数据集提供每个序列的真实位姿 TXT 文件, 根据对应两条轨迹的时间戳, 通过适当的特殊欧氏群(Special Euclidean Group, SE(3)) [18]变换来对齐两条轨迹, 具体变换矩阵  $T$  如公式 1 所示。参考 TUM 数据集, 根据相对位姿误差(Relative Pose Error, RPE)和绝对轨迹误差(Absolute Trajectory Error, ATE)对算法的性能进行评估。

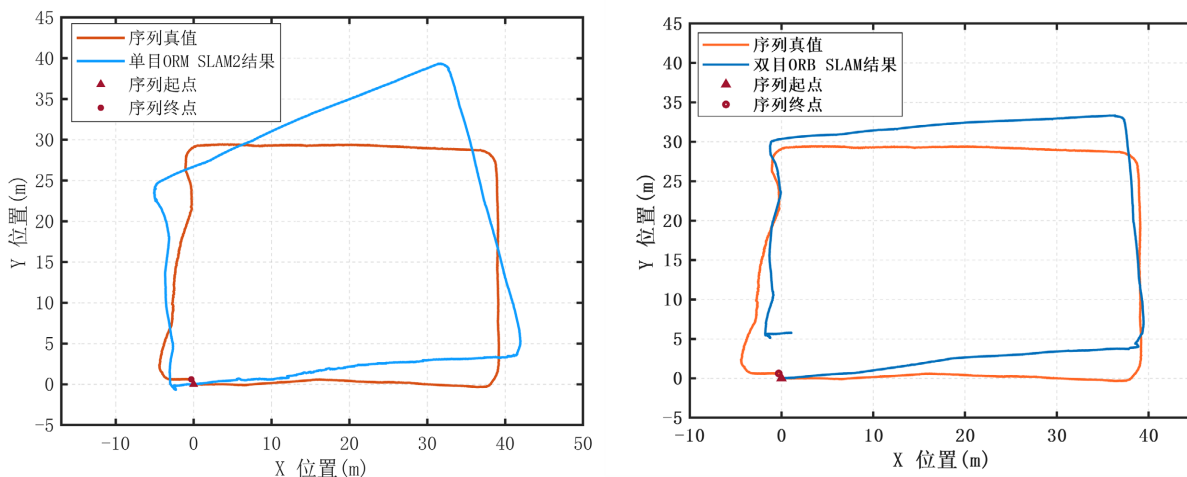
$$SE(3) = \left\{ T = \begin{bmatrix} R & t \\ O^T & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4} \mid R \in SO(3), t \in \mathbb{R}^3 \right\} \quad (1)$$

其中, 变换矩阵  $T$  左上角为旋转矩阵, 右侧为平移向量, 左下角为 0 向量, 右下角为 1, 用表示从坐标系 2 到坐标系 1 的变换。

在本数据集的两个较困难的序列 Indoor\_HARD\_001、Indoor\_HARD\_002 上评估最先进的 SLAM 系统 ORB-SLAM2 的双目和单目版本、ORB-SLAM3 的单目版本视觉定位算法的性能, 使用绝对轨迹误差 (ATE) [6]用于评估指标。算法获得的轨迹都通过 EVO 工具与地面实况对齐, 以获得 ATE 误差。

在序列 Indoor\_HARD\_001、Indoor\_HARD\_002 中, 使用小广角相机单目数据来运行的 ORB SLAM3 算法运行不佳, 出现严重的位置估计错误、定位丢失情况。虽然 ORB-SLAM3 定位算法通过自适应直方图均衡化方法解决光照不佳问题, 但算法未能在黑暗场景中提取足够的特征点, 这就导致了严重的估计错误。

在序列 Indoor\_HARD\_002 中, 使用小广角相机单目和双目数据来运行 ORB-SLAM2 算法测试效果较好, 图 12 中显示 ORB-SLAM2 的双目和单目算法在 Indoor\_HARD\_002 序列中的表现。为了更好的比较, 在运行 ORB-SLAM2 时, 禁用了算法中循环闭环部分。



**Figure 12.** Results of the visual localization algorithm run on the Indoor\_HARD\_002 sequence

**图 12.** 视觉定位算法在 Indoor\_HARD\_002 序列上运行得到的结果

在本数据集中, ORB-SLAM2 的双目版本的误差达到了米级, 而根据文献[19], 此方法在 EuRoC MAV 数据集上获得的 ATE 不大于 0.3 m。本数据集重点关注的光照频繁变化条件, 给 ORB-SLAM2 也带来了巨大的挑战。

结果表明, 现有最先进的视觉定位系统可能在现有的数据集基准中测试评估表现良好, 但它们在本数据集中至少在一种情况下表现不佳, 需要进一步努力来提高定位性能, 仍需要显著改进才能搭载在日常生活中的服务机器人上。

## 5. 结论

相比于正常光照条件下的视觉定位数据集, 供室内复杂光照环境的定位数据集的规模相对较少。为推动通过数据集构建来实现室内环境与视觉定位算法提升之间衔接的研究, 本文主要从视觉数据集的组织架构、视觉数据集构建流程和数据集的使用与评价三方面进行了探索, 并推出了用于室内复杂光照环境的定位数据集。该研究的贡献主要在于: 1) 通过构建以轮式机器人为载体的数据采集平台并结合对传感器选型与同步方法的讨论, 归纳出用于室内复杂光照环境数据集采集平台的设计思路。2) 根据采集平台的设计思路, 本文发布了一个带有地面真实轨迹的视觉定位数据集, 除了包含正常光照下的数据, 该数据集还包含强光照、弱光照和频繁变化光照等条件下的数据, 数据量超过 400 G, 时间超过 1200 s。这些数据都可用于视觉导航算法的压力测试, 可以作为目前此场景公开数据的有效补充, 并为探索机器人领域研究中其他方向提供代表性数据, 包括多传感器融合、建图和导航。3) 本文也使用自制数据集对当前主流的 SLAM 系统进行了测试, 并对实验方法和实验结果进行了讨论, 这也为视觉定位算法未来潜在的发展方向提供参考。

在未来的研究中, 研究者将继续从复杂光照条件间的特征关联方式和室内定位机理进行研究, 希望通过对数据集的不断完善, 为视觉定位算法提供一个可靠的综合评测基准。

## 基金项目

国家重点研发计划(No. 2018YFB1305003)。

## 参考文献

- [1] Geiger, A., Lenz, P. and Urtasun, R. (2012) Are We Ready for Autonomous Driving? The KITTI Vision Benchmark Suite. *IEEE Conference on Computer Vision & Pattern Recognition*, Providence, 16-21 June 2012, 3354-3361. <https://doi.org/10.1109/CVPR.2012.6248074>
- [2] Burri, M., et al. (2016) The EuRoC Micro Aerial Vehicle Datasets. *The International Journal of Robotics Research*, **35**, 1157-1163. <https://doi.org/10.1177/0278364915620033>
- [3] Smith, M., et al. (2009) The New College Vision and Laser Data Set. *The International Journal of Robotics Research*, **28**, 595-599. <https://doi.org/10.1177/0278364909103911>
- [4] Carlevaris-Bianco, N., Ushani, A.K. and Eustice, R.M. (2016) University of Michigan North Campus Long-Term Vision and Lidar Dataset. *The International Journal of Robotics Research*, **35**, 1023-1035. <https://doi.org/10.1177/0278364915614638>
- [5] Pfrommer, B., et al. (2017) PennCOSYVIO: A Challenging Visual Inertial Odometry Benchmark. 2017 *IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, 29 May-3 June 2017, 3847-3854. <https://doi.org/10.1109/ICRA.2017.7989443>
- [6] Schubert, D., et al. (2018) The TUM VI Benchmark for Evaluating Visual-Inertial Odometry. 2018 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Madrid, 1-5 October 2018, 1680-1687. <https://doi.org/10.1109/IROS.2018.8593419>
- [7] Lee, A.J., et al. (2019) ViViD: Vision for Visibility Dataset.
- [8] Huai, J., et al. (2019) Segway DRIVE Benchmark: Place Recognition and SLAM Data Collected by a Fleet of Delivery Robots.
- [9] Ramezani, M., et al. (2020) The Newer College Dataset: Handheld LiDAR, Inertial and Vision with Ground Truth. 2020 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Las Vegas, 24-30 October 2020, 4353-4360. <https://doi.org/10.1109/IROS45743.2020.9340849>
- [10] Tang, D., et al. (2016) AprilTag Array-Aided Extrinsic Calibration of Camera Laser Multi-Sensor System. *Robotics and Biomimetics*, **3**, 13. <https://doi.org/10.1186/s40638-016-0044-0>
- [11] Olson, E. (2011) AprilTag: A Robust and Flexible Visual Fiducial System. 2011 *IEEE International Conference on Robotics and Automation*, Shanghai, 9-13 May 2011, 3400-3407. <https://doi.org/10.1109/ICRA.2011.5979561>
- [12] Furgale, P.T., Rehder, J. and Siegwart, R.Y. (2013) Unified Temporal and Spatial Calibration for Multi-Sensor Systems. 2013 *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Tokyo, 3-7 November 2013, 1280-

1286. <https://doi.org/10.1109/IROS.2013.6696514>
- [13] Maye, J., Furgale, P.T. and Siegwart, R.Y. (2013) Self-Supervised Calibration for Robotic Systems. <https://doi.org/10.1109/IVS.2013.6629513>
- [14] Zhang, Q. and Pless, R. (2004) Extrinsic Calibration of a Camera and Laser Range Finder (Improves Camera Calibration). 2004 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vol. 3, 2301-2306.
- [15] Liu, S., *et al.* (2021) Brief Industry Paper: The Matter of Time a General and Efficient System for Precise Sensor Synchronization in Robotic Computing. 2021 *IEEE 27th Real-Time and Embedded Technology and Applications Symposium (RTAS)*, Nashville, 18-21 May 2021, 413-416. <https://doi.org/10.1109/RTAS52030.2021.00040>
- [16] CamLaserCalibraTool. 2018. <https://github.com/LeatherWang/CamLaserCalibraTool>
- [17] Jia, S., *et al.* (2021) A Cross-Correction LiDAR SLAM Method for High-Accuracy 2D Mapping of Problematic Scenario. *ISPRS Journal of Photogrammetry and Remote Sensing*, **171**, 367-384. <https://doi.org/10.1016/j.isprsjprs.2020.11.004>
- [18] Cadena, C., *et al.* (2016) Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age. *IEEE Transactions on Robotics*, **32**, 1309-1332. <https://doi.org/10.1109/TRO.2016.2624754>
- [19] Mur-Artal, R. and Tardós, J.D. (2017) ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. *IEEE Transactions on Robotics*, **33**, 1255-1262. <https://doi.org/10.1109/TRO.2017.2705103>