

自动机器人依赖于同时定位与地图构建 (SLAM) 来创建环境地图并在其上标出自身位置。在其中, 闭环检测是 SLAM 的关键组件, 有助于纠正漂移并提高生成地图的准确性。当机器人或自动驾驶车辆在环境中探索时, 测距及传感器测量中的微小误差会随着时间的推移而累积, 导致估算轨迹的不一致。闭环检测能检测系统何时重新访问先前已映射的位置, 允许通过优化技术 (如姿态图优化或捆绑调整) 修正过去的姿态估计。尽管闭环检测对于精确的 SLAM 至关重要, 但现有的数据集在开发强大的闭环检测算法时存在显著限制。现有数据集受限于观点多样性的不足、场景多样性的有限或闭环检测实例之间的转换元数据含糊不清。这些限制阻碍了面向各种现实场景的无缝操作之观点不变闭环检测方法的开发。LoopDB 通过一组被严格控制的同时场景的多角度捕捉、准确的转换元数据及高度的室内外场景多样性专门解决这些挑战。自动机器人开发在实验、开发和测试算法以便于真实场景中部署时, 严重依赖大量的真实世界数据。计算机视觉社区长期以来偏爱基于基准的驱动方法, 最近, 各种基于视觉的自动机器人数据集已公开提供, 例如 KITTI 数据集和 Euroc 数据集。LoopDB 采用了一种不同于现有数据集的方法, 专注于回环闭合问题, 而不是一般的 SLAM 或导航。我们的方法包含在每个场景下获取多个视图 (每个场景五个视图), 在 200 多个不同位置提供精确的相机位姿信息, 使研究人员能够在受控但具有挑战性的视角变化下测试和评估回环闭合算法。与轨迹相关的数据集会随机发生回环闭合不同, LoopDB 提供了具有已知变换的保证回环闭合, 允许对检测准确性和几何一致性进行量化评估。这使得可以开发能够处理显著视点变化的基准回环闭合算法, 这是现实世界机器人应用中常见的问题, 而现有数据集往往没有很好地解决这些问题。

在表中展示了 LoopDB 和计算机视觉研究中著名的数据集之间的对比。虽然 LoopDB 较小 (1000 张图像), 但它包含了更多样化的场景, 与 KITTI、Oxford RobotCar 或 EuRoC MAV 数据集不同。回环闭合检测可以最小化定位漂移并提高地图绘制精度。尽管以下数据集并非为了回环闭合的基准测试而设计, 但由于缺乏更好的替代方案, 这些数据集被使用。KITTI 数据集是计算机视觉和机器人社区中常用的数据集。该数据集包括在城市环境和农村道路中驾驶车辆收集的立体相机图像、LiDAR 扫描和 GPS 测量。尽管 KITTI 并非为回环闭合检测设计, 但它对 SLAM 和视觉里程计算研究作出了重大贡献。4Seasons 数据集 [?] 提供了全年在苏黎世多种天气条件下的数据。4Seasons 的特别之处在于系统地覆盖了具有挑战性的环境条件, 如雪、雨、阳光, 以及同一路线上随季节变化的树叶变化, 以在不同外观条件下评估视觉位置识别和回环检测算法。KITTI-360 [?] 通过提供城市场景的密集 360° 语义和几何注释

节。我们使用齐次坐标表示图像中的点, 其表示法和其他地方标准一致。这在处理平移、旋转和缩放等变换时非常重要, 因为它允许我们通过矩阵乘法来表达这些操作。例如, 在 2D 空间中, 一个点 X 表示为 $X = [X \ Y \ 1]^T$, 其中 X 和 Y 是图像平面中点的坐标, 最后一个分量设置为 1 以便于仿射变换。在这些点上操作的矩阵被写为单应性矩阵 H , 它编码了视图之间的变换。该数据集的关键特征是每张图像的变换数据。由于此数据集是为闭环检测和 SLAM 系统设计的, 对于任何场景的不同视图之间准确的变换数据是迫切需要的。每对图像都有一个旋转矩阵和一个平移向量分配, 以避免用户自行计算图像之间的相对姿态。由单应性捕获的两个图像之间的变换, 记为 H , 是一个 3×3 矩阵, 包含了旋转和平移部分。从第一幅图像坐标 X_1 到第二幅图像坐标 X_2 的变换如下所示:

$$X_2 = H \cdot X_1 \quad (1)$$

, 其中 X_1 和 X_2 表示齐次坐标系中的点:

$$X_1 = \begin{bmatrix} X_1 \\ Y_1 \\ 1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} X_2 \\ Y_2 \\ 1 \end{bmatrix} \quad (2)$$

单应矩阵 H 包含两个部分: 旋转矩阵 R 和平移向量 t 。在 3D 变换的情况下, 矩阵形式如下:

$$H = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \quad (3)$$

, 其中 R 是 3×3 旋转矩阵, t 是 3×1 平移向量, 描述了两个图像之间相机的移动。图像间的平移由向量 $t = [t_x, t_y, t_z]$ 表示。我们使用四元数来代替直接存储旋转矩阵, 因为它们更紧凑且可以避免万向节锁的问题。一个四元数 q 表示为 $q = [q_x, q_y, q_z, q_w]$, 其中 q_x, q_y, q_z 是四元数的虚部, q_w 是标量部分。四元数用于表示相机的 3D 方向, 并且它们可以实现不同方向之间的平滑过渡。两个图像之间的完整变换表示为:

$$\begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} = R \cdot \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} + t \quad (4)$$

A. 数据格式

LoopDB 包含图像和 CSV 元数据文件，并组织成两个数据流。每个场景（子图）由五幅具有不同视角的图像组成，从具有零旋转和位移的根图像开始。数据集结构如图 ?? 所示，包含以下内容。

这种组织方式使得在图像之间进行空间关系的准确重建成为可能，支持姿态估计、视觉里程计和 SLAM 应用。

如上所述，LoopDB 使用两个大型 CSV 格式的元数据文件来对每个图像进行详细标注。CSV 文件提供了每个图像对结构的旋转和位移信息。所有行都映射到对应图像时间戳的图像上，并伴有特定的旋转（四元数）、位移或其他元数据。

数据集及所有必要的脚本可以在 GitHub 上公开获取（<https://github.com/RovisLab/LoopDB>），而数据集图像可以在 Zenodo 上获取（www.doi.org/10.5281/zenodo.15201910）。研究人员可以直接下载图像文件和 CSV 元数据，或使用提供的 Python 脚本将其集成到他们的 SLAM/回环检测管道中。库中包含详细的文档，提供逐步的设置说明、实验指南和定制选项。

为确保高质量的数据处理，使用了一系列软件库和编程语言。最显著的是：

这些工具因其在计算机视觉和机器人学中的灵活性和广泛使用而被选择，确保研究社区的易用性。

LoopDB 解决了回环检测中特定的挑战，这对于稳健的 SLAM 系统至关重要，支持传统的基于特征的方法和新兴深度学习方法。研究人员可以多种方式使用 LoopDB。对传统方法的基准测试：该数据集为在具有挑战性的视点变化下评估基于特征的算法（SIFT [?], ORB [?], BRIEF [?], DAISY [?]) 提供了一个标准化的测试。学习型方法的训练数据：LoopDB 拥有跨越 200 多个不同场景的 1000 多张高质量图像，为训练深度学习模型提供了足够的多样性，精确的变换元数据有助于正样本和负样本对的创建。基于 Transformer 的方法：该数据集非常适合评估利用注意力机制在视点变化中建立稳健对应关系的 Transformer 架构。像 MixVPR [?] 和 TransVPR [?] 等方法可以通过使用 LoopDB 的多角度捕获进行有效的基准测试。高级应用：LoopDB 支持使用 RANSAC 变体或基于学习的方法（如 SuperGlue [?]）进行几何验证研究，并能够评估现代 SLAM 系统如 DROID-SLAM [?] 和 DeepFactors [?]。该数

据集的受控环境允许利用精确的变换元数据隔离并评估闭环组件。通过提供专注于受控视点变化的专用数据集，LoopDB 在研究生态系统中弥补了关键空白，促进了对这一基础 SLAM 挑战的传统和学习型方法的有针对性的发展和评估。

数据是使用 48MP 华为 Nova 7i 相机拍摄的，选择该相机是因为其在多样的光照和环境条件下捕获高质量图像的能力。图像在不同比例的 ISO 设置和曝光时间下拍摄，以引入亮度、对比度和噪声变化，这对在挑战性条件下测试算法至关重要。在没有 GPS 定位的情况下，我们依赖于仔细的相机布局来从多个视角捕获每个场景。计算并存储了连续图像之间的变换矩阵于元数据文件中。如图 ?? 所示，每个场景都从多个视点拍摄，同时保持一致的中心参考，确保图像序列的一致性，以研究跨视角的视觉连续性。相机经过焦距、主点和镜头畸变的校准，以确保几何精确的图像，以进行可靠的特征匹配。通过精确手动控制相机位置建立了真实对齐，确保每个场景从不同视角拍摄，同时在序列中的所有五张图像中保持一致的焦点。

B. 校准与同步

校准对于校正畸变和确保准确的相机属性至关重要。在该数据集中，相机经过校准焦距、主点和镜头畸变，确保了几何精确的图像，在不同视角之间特征可以可靠匹配。由于数据采集使用的是单个相机，图像被手动同步以确保一致的对齐。该方法保证了连续帧保持一致的时间关系，在它们之间变形最小。

C. 真实对齐

真实对齐通过拍摄时精确手动控制相机位置来建立。每个场景都故意从不同的视角拍摄，同时在一个系列中的所有五张图像中保持一致的焦点。虽然这种真实对齐是基于视觉对齐而不是外部位置测量，但它非常适用于 SLAM 应用，尤其是那些不需要外部定位数据的单目视觉系统。

I. 实验评估

为了验证 LoopDB，我们使用四个描述子（SIFT [?], ORB [?], DAISY [?] 和 BRIEF [?]) 进行了特征对应实验，以评估投影误差。主要发现包括：距离效应：误差随着距离参考图像的增加而增加。空间模式：误差显示出一致的空间相关模式。纹理重要性：纹理良好的区域产生更好的特征对应。视点挑战：显著的误差（在某些情况下为 > 100 像素）显示了数据

集的复杂性。我们评估了数据集中转换的平均重投影误差 (见图 ??)。如表 I 所示, BRIEF 表现最佳 (平均误差 7.47 像素), 其次是 DAISY (14.76 像素), 而 ORB 和 SIFT 误差较高 (分别为 30.66 像素和 35.95 像素)。1→2 的转换最具挑战性, 误差范围从 14.64 像素 (BRIEF) 到 163.51 像素 (SIFT)。表 I 确认了 BRIEF 在所有转换中的优越表现。

我们还使用相同的描述子将 LoopDB 与五个流行的数据集进行了比较 (表 II)。LoopDB 展示了三个关键优势: 高特征密度——显著提取了更多特征 (SIFT: 810, BRIEF: 679, DAISY: 688); 检测器的多功能性, 无论是 SIFT (0.78 像素) 还是二进制描述子表现都很好; 以及几何一致性, 尽管与其他数据集相比提取了更多特征, 但仍保持了具竞争力的误差率。这些结果确立了 LoopDB 作为环闭合检测方法的有效基准, 其五帧链结构使得能够评估算法在连续转换中的性能, 模拟实际 SLAM 应用中的误差累积。

TABLE I
跨过渡和方法的重投影误差。

Method	Transition Pairs (Error/Match)				Overall	
	0→1	1→2	2→3	3→4	Avg	Match
SIFT	4.16 731	163.5 454	8.35 591	3.71 1052	35.9	163.5
ORB	10.2 679	121.8 767	12.4 777	9.0 958	30.7	121.8
Daisy	1.42 490	65.3 281	5.30 436	1.82 847	14.8	65.3
BRIEF	8.39 770	14.6 690	8.71 723	5.60 1025	7.5	14.6

我们还使用相同的描述子将 LoopDB 与五个流行的数据集进行了比较 (表 II)。LoopDB 展示了三个关键优势: 高特征密度——显著提取了更多特征 (SIFT: 810, BRIEF: 679, DAISY: 688); 检测器的多功能性, 无论是 SIFT (0.78 像素) 还是二进制描述子表现都很好; 以及几何一致性, 尽管与其他数据集相比提取了更多特征, 但仍保持了具竞争力的误差率。这些结果确立了 LoopDB 作为环闭合检测方法的有效基准, 其五帧链结构使得能够评估算法在连续转换中的性能, 模拟实际 SLAM 应用中的误差累积。

TABLE II
数据集间平均重投影误差比较 (像素)

Dataset	SIFT	ORB	DAISY	BRIEF
LoopDB (Ours)	0.78	44.67	5.64	4.46
KITTI [?]	1.16	33.87	1.24	1.82
Oxford [?]	1.32	33.70	5.47	10.64
EuRoC [?]	0.56	0.53	0.79	0.79
4Seasons [?]	10.28	35.42	0.81	0.93
KITTI-360 [?]	8.62	100.58	1.11	1.97

II. 结论

本研究介绍了一个具有挑战性和多样性的数据集, 专门用于回环闭合检测和 SLAM。该数据集包含了超过 1000 张高质量的图像, 这些图像在各种环境中拍摄, 如公园、健身房和停车场。它还提供每个场景的多个视角, 以及对应的变换数据, 使其成为训练和测试自主导航算法的宝贵资源。未来的增强可能包括集成附加传感器, 如 GPS 和 IMU, 以改善姿态估计和地面真实对齐。此外, 引入动态元素, 如移动物体和变化的天气条件, 将进一步增加数据集的逼真性。这些改进将允许在高度动态的环境中测试 SLAM 和回环闭合检测算法, 这是迈向开发稳健和适应性强的自主系统的关键步骤。

我们计划扩展 LoopDB, 以使数据集对传统方法和基于学习的方法都更有用。我们的策略包括: 环境多样性: 我们将添加更多具有挑战性的位置, 例如繁忙的城市区域、工业园区和地下结构。条件变化: 我们将在雨、雪、雾等具有挑战性的天气条件下捕捉序列, 具有变化的照明 (黎明、黄昏、夜晚) 和季节变化。动态元素: 包括移动物体的序列和不同人群规模。轨迹复杂性: 加入具有复杂轨迹的较长序列, 包括在单个序列内的多个不同大小的回环闭合。扩展后的数据集将支持深度学习应用, 提供超过 7000 张图像, 以实现神经网络对回环闭合检测的有效训练。