Journal of Intelligent & Robotic Systems (2024) 110:146
https://doi.org/10.1007/s10846-024-02172-6

REGULAR PAPER

MULO: LiDAR Odometry via MUlti-Landmarks Joint Optimization

Jun Liu[1] · Zhengnan He[1] · Xiaoyu Zhao[1] · Jun Hu[2] · Shuai Cheng[2] · Wei Liu[1]

Received: 11 July 2023 / Accepted: 22 August 2024

Abstract
At the present stage, LiDAR-based SLAM solutions are dominated by ICP and its variants, while the BA optimization method that can improve the pose consistency has received little attention. Therefore, we propose MULO, a low-drift and robust LiDAR odometry using BA optimization with plane and cylinder landmarks. In the front-end, a coarse-to-fine direct pose estimation method provides the prior pose to the back-end. And in the back-end, we propose a novel three-stage landmark extraction and data association strategy for plane and cylinder, which is robust and efficient. Meanwhile, a stable minimum parameterization method for cylinder landmarks is proposed for optimization. In order to fully utilize the LiDAR information at long distances, we propose a new sliding window structure consisting of a TinyWindow and a SuperWindow. Finally, we jointly optimize the two kinds of landmarks and scan poses in this sliding window. The proposed system is evaluated on public dataset and our dataset, and experimental results show that our system is competitive compared with the state-of-the-art LiDAR odometrys.

Keywords SLAM · LiDAR odometry · MUlti-Landmarks · Plane and cylinder landmarks

1 Introduction

Simultaneous localization and mapping (SLAM) plays an important role in automatic driving and other tasks involved in intelligent mobile robot. In recent years, SLAM schemes based on various sensors have been proposed. Among them, the LiDAR-based method can be applied in most scenarios because it is not affected by lighting conditions. Nevertheless, as a well-recognized method that can improve the consistency of pose, bundle adjustment(BA) has rarely been investigated in the LiDAR-based system. This paper devotes to exploring the LiDAR BA method that are applicable to indoor and outdoor environments.

Nowadays, LiDAR-based methods are mainly based on the ICP framework [1] and its variants [2–5], which realize the pose estimation through a scan-to-map or scan-to-model registration. However, the sparsity and disorder of the LIDAR point cloud make it impossible to be observed repeatedly in multiple frames directly. Although some feature extraction methods for point cloud have been proposed [6–8], the dense point cloud structure requirement makes these methods unsuitable for sparse LiDAR points.

As one of the few landmarks that can be observed repeatedly in multiple frames with sparse point clouds, the plane landmark is receiving increasing attention from the LiDAR SLAM systems [9–15]. However, these systems are only applicable to the structured scenes with enough plane landmarks and are not highly adaptable in outdoor environments.

In this paper, the BA method for LiDAR is further studied. We find that using plane landmarks alone cannot cope with the environment lacking planes, while there is a large amount of cylinder information such as tree trunks and light

Jun Liu and Zhengnan He contributed equally to this work.

✉ Wei Liu
lwei@reachauto.com

Jun Liu
2010354@stu.neu.edu.cn

Zhengnan He
2270917@stu.neu.edu.cn

Xiaoyu Zhao
928674428@qq.com

Jun Hu
hu.jun@reachauto.com

Shuai Cheng
cheng.shuai@reachauto.com

[1] School of Information Science and Engineering, Northeastern University, NO. 3-11, Wenhua Road, Shenyang 110819, P. R. China

[2] Neusoft Reach Automotive Technology Company, NO.2 Xin Xiu Street A4, Shengyang 110179, P. R. China

Published online: 04 October 2024

poles, especially in urban roads scene. Therefore, we propose a novel LiDAR odometry MULO(as shown in Fig. 1, which combines the plane and the cylinder landmarks for BA optimization.

The main contributions of this paper are as follows:

- A new robust and efficient three-stage landmark extraction and data association strategy for plane and cylinder.
- A new calculating stable parameterization method for cylinder landmark.
- A novel sliding window structure consisting of a Tiny-Window and a SuperWindow that can fully utilize long-range information from LIDAR in the BA optimization stage.
- Extensive experiments are conducted on UrbanLoco dataset [16] and our dataset with outdoor and indoor scenarios. Our system MULO is competitive compared with the state-of-the-art LiDAR odometrys.

2 Related Work

For the current LiDAR-based SLAM system, the ICP and its variants are still the main methods for pose estimation. LOAM [2] extracts plane and edge points by calculating the curvature on each scan-line, then a scan-to-scan and a scan-to-map registration are conducted successively by minimizing the point-to-line(plane) distance. LEGO-LOAM [3] is a ground-optimized LiDAR odometry for unmanned ground vehicles, which segments the ground points to assist feature extraction. A two-step Levenberg-Marquardt optimization is conducted to perform real-time six degree-of-freedom pose estimation. F-LOAM [4] is another computationally efficient and accurate version of LOAM, which use a non-iterative two-stage distortion compensation method to reduce the computational cost. MULLS [5] is a versatile LiDAR SLAM that jointly optimizes point-to-point (plane, line) error metrics with a linear approximation to estimate the ego-motion. In the back-end, MULLS conducts a hierarchical pose graph optimization to reduce the drift of trajectory. All the above systems use the ICP variant methods to perform pose estimation by extracting different features. However, these features can not be repeatedly observed in multiple frames.

For the purpose of extracting repeatable observed features in multiple frames, several feature-based methods have been explored. Fast Point Feature Histograms (FPFH) [7] forms a multi-dimensional histogram to geometrically describe the nearest neighbors of points, which can be used in real-time applications. Viewpoint Feature Histogram (VFH) [8], which proposes a new descriptor consisting of an extended FPFH descriptor and a new signature, achieves better performance than FPFH. However, for these methods, the requirement of dense point clouds makes them unsuitable for sparse LiDAR points.

In recent years, some SLAM systems have begun to consider plane landmarks. LIPS [10] and Eigen-Factors [11] use different parameterization methods to add plane landmarks into the optimization process. LIC-FUSION2.0 [17] and MINS [18] add plane landmarks into the multi-sensor optimization process. BALM [9] introduces a new kind of BA method using plane and edge landmarks. Instead of optimizing the landmark parameters and scan poses simultaneously, BALM analytically solves landmarks in closed-form before the BA optimization with a second order derivative and a highly efficient Gauss-Newton method. π-LSAM [15] proposes a plane adjustment(PA) method suitable for indoor environments, which can reduce the optimization complexity by π-factor. However, in the scene lack of plane landmarks, these systems are easy to degenerate, and it is not accurate to directly treat trunks and light poles as line landmarks.

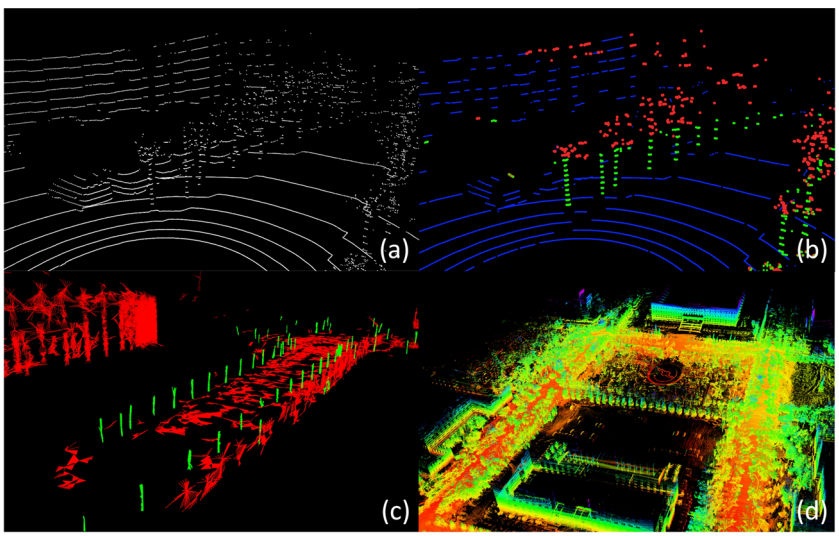

Fig. 1 An illustration of the proposed MULO system: (a) raw LiDAR points, (b) different types of points extracted from the scan(noise, cylinder, plane), (c) plane landmarks and cylinder landmarks extracted from the local map(plane, cylinder), (d) map generated by MULO

Furthermore, they are not suitable for complex outdoor environments without considering the dynamic objects. In this paper, cylinder landmarks are proposed to avoid degeneration, and a foreground segmentation method is used to eliminate the influence of dynamic objects.

3 System Overview

The pipeline of the proposed system is shown in Fig. 2. The system receives input from a 3D LiDAR, the ego-motion is calculated by a coarse-to-fine direct pose estimation method. Meanwhile, the planar points and the columnar points are extracted from the raw points. Then a foreground segmentation is performed to filter the dynamic objects for the planar points, and motion compensation is directly conducted on these two kinds of feature points. With the cursory pose calculated by the front-end, data association is performed for both types of landmarks. Then new landmarks are extracted from the feature points with no corresponding landmarks. At last, a maximum a posteriori (MAP) problem is solved to jointly optimize the scan poses and the parameterized cylinder and plane landmarks in a sliding window.

4 Coarse-to-fine Direct Pose Estimation Front-end

In the front-end, we refer to the direct method proposed by FAST-LIO2 [19]. Since only LiDAR sensor is used in our system, the LiDAR raw point clouds are directly downsampled as different resolutions. The data association and a scan-to-map direct registration are performed from low resolution to high resolution with an initial pose predicted by the uniform motion model. We minimize the point-to-plane distance by the L-M method to obtain the pose of the current frame:

$$\min_{\boldsymbol{T}} \sum_{k=1}^{N} \left\| \boldsymbol{n}_k^T \left(\boldsymbol{T}\boldsymbol{p}_k - \boldsymbol{c}_k \right) \right\|^2, \tag{1}$$

where $\boldsymbol{T} \in SE(3)$ denotes the pose of the current frame, $\boldsymbol{p}_k$ denotes the point in the current LiDAR scan described in the current frame. $\boldsymbol{n}_k$ and $\boldsymbol{c}_k$ denote the Hessen normal and the center of the corresponding planar patch extracted in the global point cloud map. Finally, motion compensation is performed using the uniform motion model.

5 Landmark Parameterization

5.1 Plane Landmark

We use plane patch [18] consisting of a center point $\boldsymbol{p}_{pp}$ and a Hesse normal $\boldsymbol{n}_{pp}$ to describe a plane landmark. Although this representation is convenient for data association, it is over-parameterized for non-linear optimization. Accordingly, closest point(CP) [10] is used as the minimal representation of the plane landmark in the optimization stage. The CP representation indicating the point on the plane with minimum distance to the measured frame's origin can be described as a Hesse normal vector $\boldsymbol{n}$ and a distance scalar d. It has the following relations:

$$\boldsymbol{\Pi} = \boldsymbol{n}d, \tag{2}$$

$$\begin{bmatrix} \boldsymbol{n} \\ d \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Pi} / \|\boldsymbol{\Pi}\| \\ \|\boldsymbol{\Pi}\| \end{bmatrix}. \tag{3}$$

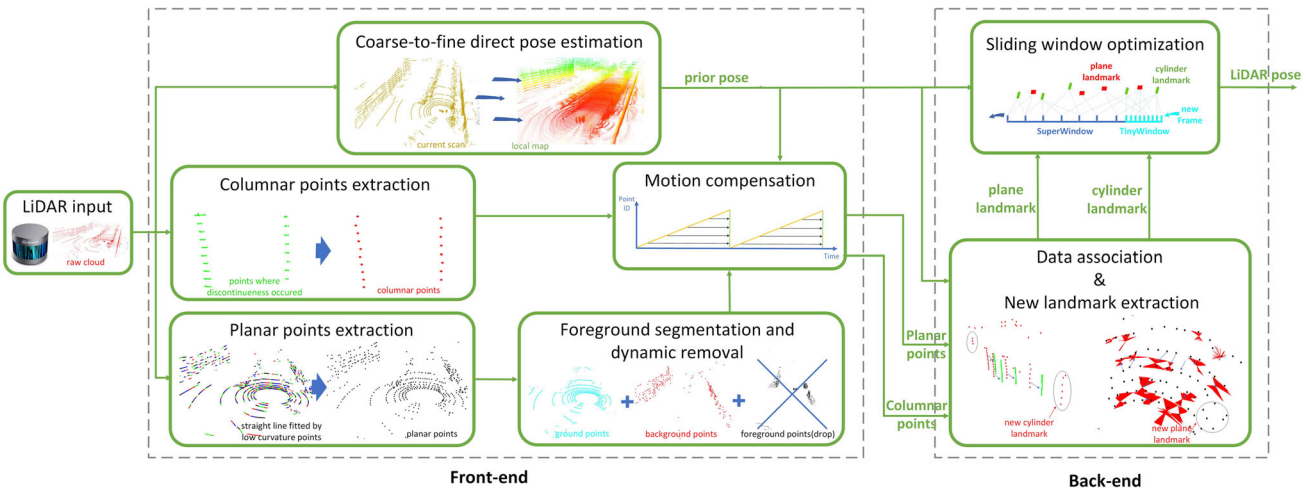

Fig. 2 System overview. The system consists of a front-end and a back-end. Front-end performs direct pose estimation and feature points extraction. Back-end includes data association, landmark extraction and sliding window optimization

The transformable relationship between the plane patch and the CP is as follows:

$$\boldsymbol{\Pi} = \boldsymbol{n}_{pp}(\boldsymbol{n}_{pp}{}^{T}\boldsymbol{p}_{pp}). \tag{4}$$

5.2 Cylinder Landmark

We propose a set of novel representations to describe the cylinder landmark. A center point $\boldsymbol{p}_{cc}$, a direction normal $\boldsymbol{n}_{cc}$ and a radius r_{cc} describe the cylinder landmark for data association, we call them cylinder chunk(CC). In addition, a minimal parameterization form called minimal cylinder(MC) is proposed for optimization, which use a 3D unit vector $\boldsymbol{n}_{mc}$ with 2 DOF, a radius r_{mc} with 1 DOF and a 3D position $\boldsymbol{p}_{mc}$ with 2 DOF to describe the 5 DOF cylinder landmark. In this method, $\boldsymbol{p}_{mc}$ is constrained on a plane whose Hesse normal directs the axis where $\boldsymbol{n}_{cc}$ has the maximum component. An illustration of this parameterization method is shown in Fig. 3. In this way, point to cylinder distance can be directly calculated without other transformations. The transformable relationship between the CC and the MC is as follows:

$$\begin{bmatrix} \boldsymbol{n}_{mc} \\ \boldsymbol{p}_{mc} \\ r_{mc} \end{bmatrix} = \begin{bmatrix} \boldsymbol{n}_{cc}/\left\|\boldsymbol{n}_{cc}\right\| \\ \boldsymbol{p}_{cc} \\ r_{cc} \end{bmatrix}. \tag{5}$$

The coordinate components of $\boldsymbol{p}_{mc}$ along the axis corresponding to the largest coordinate component of $\boldsymbol{n}_{cc}$ are fixed during the optimization process.

6 Landmark Extraction And Data Association

For the the plane landmark and the cylinder landmark, a three-stage flow for landmark extraction and data association is proposed, which are feature points extraction, observation generation and new landmark extraction.

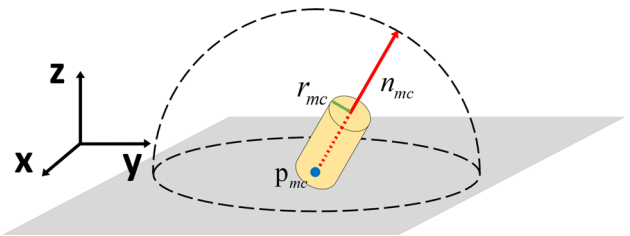

Fig. 3 An example illustrates the MC. The red line denotes the 2 DOF unit vector $\boldsymbol{n}_{cc}$. Since the maximum component of $\boldsymbol{n}_{cc}$ is along the z-axis, the blue point $\boldsymbol{p}_{mc}$ is constrained on a grey plane whose Hesse normal directs the z-axis. In addition, the green line denotes the radius r_{mc}

6.1 Cylinder Landmark

As shown in Fig. 4, we first detect the points where discontinuousness occurred in each consecutive scan-line. If the distance between two points where discontinuousness occurred is close, the points between these two points are regarded as the candidate feature points on a cylinder landmark. Finally, we use the mean value point called columnar point to represent this group of points to accelerate the data association process. A detailed flow of this approach is shown in Algorithm 1.

Algorithm 1 Columnar points extraction.

```
for all scanline_i ∈ LiDAR scan do
  bFindSuccess ← false
  bNewCyl ← false
  for all p_{j-1}, p_j ∈ scanline_i do
    rangeDis ← Range(p_{j-1}) − Range(p_j)
    if rangeDis > MaxRangeThres then
      bNewCyl ← true
      idxStart ← j
    end if
    if bNewCyl == true then
      if −rangeDis > MaxRangeThres then
        bNewCyl ← false
        idxEnd ← j − 1
        groupWidth ← Distance(p_idxEnd, p_idxStart)
        if groupWidth < MaxWidthThres then
          bFindSuccess ← true
        end if
      end if
    end if
    if bFindSuccess == true then
      bFindSuccess ← false
      newGroup ← points between p_idxStart , p_idxEnd
      newColpoint ← MeanValue(newGroup)
      Add newGroup to vectorGroups.
      Add newColpoint to vectorColpoints.
    end if
  end for
end for
```

All cylinder landmarks in the sliding window are collected to build a landmark map in the observation generation stage. For every columnar point $\boldsymbol{p}_{col}$ in the current frame, we use a KD-tree to search the nearest cylinder landmark in the landmark map. A point to cylinder distance score is calculated to verify the correspondence between the columnar point and the cylinder landmark. The score is defined as follows:

$$score_{cyl} = \left\|\boldsymbol{n}_{cc} \times (\boldsymbol{p}_{col} - \boldsymbol{p}_{cc})\right\|. \tag{6}$$

The group of points represented by columnar point is regarded as this cylinder landmark's observation in the current frame if the above score is smaller than a threshold.

In the new landmark extraction stage, we first attempt to extract landmarks in a single frame. All columnar points

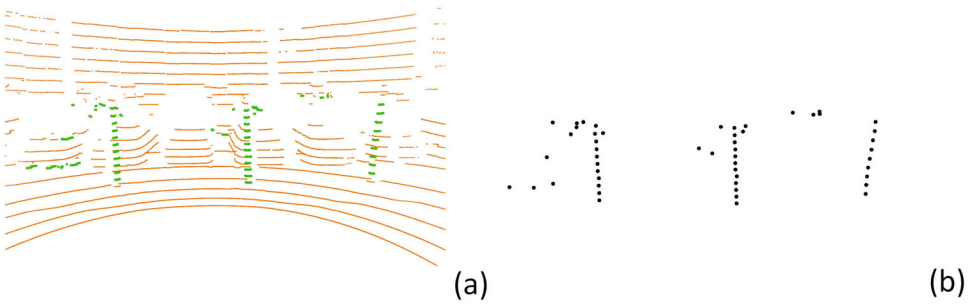

Fig. 4 Columnar points extraction. (a) Points between the two points where discontinuousness occurred are collected. (b) Columnar points are extracted from each group of points

that failed to find the corresponding landmark are organized as a KD-tree. The nearest k points to each point are collected, and a principal component analysis(PCA) method is used to fit and generate new cylinder landmarks. The criterion of the successful fitting is $\lambda 1/\lambda 2 > 10$, where $\lambda 1$ and $\lambda 2$ are the maximum and medium eigenvalues calculated by PCA. However, the sparsity of the LiDAR point cloud causes some landmarks cannot be extracted from a single frame. Therefore, we further register the remaining columnar points without corresponding landmarks in each frame into a local map using the optimized pose, then we organize the map as a KD-tree for searching and extracting landmarks like the method in the single frame. With this approach, we can perform landmark extraction and data association steadily and without omission.

6.2 Plane Landmark

As shown in Fig. 5, we find that points on a planar surface patch in a scan-line can be approximately expressed as a straight line. Thus, if the consecutive points with low curvature in each scan-line can be fitted into a straight line, we use the center point with direction vector information to represent this group of points. This point is called the planar point. With the above method, we can obtain the exact plane candidate points and significantly reduce the search time of the data association process.

A flow similar to the cylinder landmark is applied in the observation generation stage. Furthermore, a point-to-plane distance score and a direction score verify the correspondence between planar point and plane landmark. These two scores are defined as:

$$score1_{pl} = (\boldsymbol{p}_p - \boldsymbol{p}_{pp})^T \boldsymbol{n}_{pp} / \left\| \boldsymbol{n}_{pp} \right\|, \tag{7}$$

$$score2_{pl} = \boldsymbol{n}_p^T \boldsymbol{n}_{pp} / (\left\| \boldsymbol{n}_p \right\| \left\| \boldsymbol{n}_{pp} \right\|). \tag{8}$$

Similar to the new landmark extraction stage of the cylinder landmark, we extract the plane landmarks in a single frame and in local map successively. In each process, we find the nearest k planar points of each point in the planar point map, and then there are two possible situations:

- If the direction vectors of two of these points are approximately intersecting and these two points with direction information fit on a plane, a new plane landmark is generated
- If the direction vectors of three of these points are approximately parallel, we use two of these points with their direction information to fit a plane. A new plane landmark is generated if the third point with direction information is on this plane.

6.3 Dynamic Removal

Since the dynamic object is hardly detected on the cylinder landmark, we only filter the dynamic objects on the plane landmark. We used the method in [3] to segment ground points. For non-ground points, we segment the foreground and background of the scene. Background points generally appear on the wall of high buildings, and foreground points are considered dynamic objects, so the plane land-

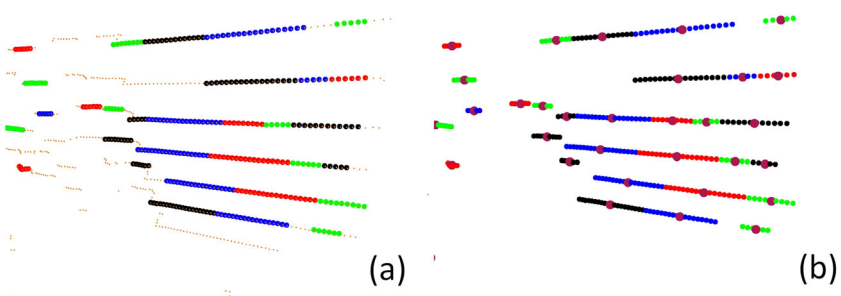

Fig. 5 Planar points extraction. (a) Points belonging to the same straight line are marked as the same color. (b) Planar points with direction information are extracted from each group of points

Table 1 Ablation experiment

Algorithm	Data sequence	
	Outdoor_1(translation error(m))	HK-Data20190316-2(RMSE w.r.t SPAN-CPT(m))
MULO	1.451	1.370
MULO-Undistort	11.750	5.421
MULO-Cylinder	Fail	1.382
MULO-SuperWindow	8.588	15.314
MULO-BA	4.276	14.258

marks existing in the foreground are removed as dynamic objects directly.

7 Sliding Window Optimization

7.1 Sliding Window Structure

Since LiDAR has a long measurement distance and a large observation perspective, it is difficult for a short sliding window used in the vision-based system to fully use the LiDAR information in a large scene. We propose a novel sliding window structure with a distance-based keyframe selection strategy. It consists of two sliding windows called TinyWindow and SuperWindow, which use 1 meter and 5 meters as the keyframe selection threshold, respectively. There are five operations in this sliding window structure:

- A new frame is added to the TinyWindow.
- The second new frame is slided out and discarded if it is not a keyframe of the TinyWindow.
- The oldest keyframe in the TinyWindow is slided out and discarded if the TinyWindow is full and it is not the keyframe of the SuperWindow.
- The oldest keyframe in the TinyWindow is slided out and added to the SuperWindow if the TinyWindow is full and it is the keyframe of the SuperWindow.
- The oldest keyframe in the SuperWindow is slided out and discarded if the SuperWindow is full.

7.2 Joint Optimization

The whole state in the sliding window is as follows:

$$\begin{cases} \boldsymbol{\chi} = [\boldsymbol{x}_0, \cdots, \boldsymbol{x}_r, \boldsymbol{c}_0, \cdots, \boldsymbol{c}_m, \boldsymbol{\Pi}_0, \cdots, \boldsymbol{\Pi}_n] \\ \boldsymbol{x}_i = [{}^w\boldsymbol{p}_i, \ {}^w\boldsymbol{q}_i], i \in [0, r] \\ \boldsymbol{c}_k = [\boldsymbol{n}_{mc_k}, \boldsymbol{p}_{mc_k}, r_{mc_k}], k \in [0, m] \end{cases}, \tag{9}$$

where $\boldsymbol{x}_i$ is the i-th LiDAR frame state in the sliding window, $\boldsymbol{c}_k$ and $\boldsymbol{\Pi}_s$ are respectively the k-th cylinder landmark parameter and the s-th plane landmark parameter expressed in their anchor frame in the sliding window.

The sum of the norm of cylinder and plane measurement residuals are used to obtain a maximum a posteriori estimation:

$$\min_{\boldsymbol{\chi}} \left\{ \sum \left\| \boldsymbol{r_c}(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{c}_k) \right\|^2 + \sum \left\| \boldsymbol{r_p}(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{\Pi}_s) \right\|^2 \right\}. \tag{10}$$

$\boldsymbol{r_c}(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{c}_k)$ is the residual of cylinder landmark measurement, and it is defined as:

$$\boldsymbol{r_c}(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{c}_k) = \left\| {}^i\boldsymbol{n}^i_{mc_k} \times ({}^i\boldsymbol{p}^j_{c_u} - {}^i\boldsymbol{p}^i_{mc_k}) \right\| - \left\| {}^i\boldsymbol{n}^i_{mc_k} \right\| r^i_{mc_k}, \tag{11}$$

$${}^i\boldsymbol{p}^j_{c_u} = {}^w\boldsymbol{q}_i{}^{-1}({}^w\boldsymbol{q}_j \, {}^j\boldsymbol{p}^j_{c_u} + {}^w\boldsymbol{p}_j) - {}^w\boldsymbol{p}_i), \tag{12}$$

where ${}^i\boldsymbol{p}^j_{c_u}$ is the u-th cylinder point of the j-th frame represented in the i-th frame, and it is the observation of the k-th cylinder landmark of the i-th frame. ${}^i\boldsymbol{n}^i_{mc_k}$, ${}^i\boldsymbol{p}^i_{mc_k}$ and $r^i_{mc_k}$ are the components of the k-th cylinder landmark $\boldsymbol{c}_k$ of the i-th frame represented in the i-th frame.

Table 2 RMSE translation error w.r.t SPAN-CPT. (meters)

Algorithm	Data sequence			
	HK-Data20190426-2	HK-Data20190316-2	HK-Data20190117	HK-Data20190426-1
LEGO-LOAM	3.431	1.473	1.427	2.859
F-LOAM	3.553	2.223	4.749	3.295
BALM	26.274	2.631	35.535	39.648
MULLS	2.507	3.017	1.282	**2.574**
MULO	**2.287**	**1.370**	**0.964**	3.016

Table 3 End-to-end translation error. (meters)

Algorithm	Data sequence		
	Indoor_1	Indoor_2	Outdoor_1
LEGO-LOAM	0.066	0.450	2.300
F-LOAM	0.189	0.596	8.520
BALM	0.061	11.268	13.996
MULLS	33.081	9.935	27.608
MULO	**0.059**	**0.338**	**1.451**

$\boldsymbol{r_p}(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{\Pi}_s)$ is the residual of plane landmark measurement, and it is defined as:

$$\boldsymbol{r_p}(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{\Pi}_s) = {}^i\boldsymbol{n}^i_{pp_s}{}^T({}^i\boldsymbol{p}^j_{p_v} - {}^i\boldsymbol{p}^i_{pp_s}) / \left\| {}^i\boldsymbol{n}^i_{pp_s} \right\|, \quad (13)$$

$${}^i\boldsymbol{p}^j_{p_v} = {}^w\boldsymbol{q}_i{}^{-1}({}^w\boldsymbol{q}_j\ {}^j\boldsymbol{p}^j_{p_v} + {}^w\boldsymbol{p}_j) - {}^w\boldsymbol{p}_i), \quad (14)$$

where ${}^i\boldsymbol{p}^j_{p_v}$ is the v-th plane point of the j-th frame represented in the i-th frame, and it is the observation of the s-th plane landmark of the i-th frame. ${}^i\boldsymbol{n}^i_{pp_s}$ and ${}^i\boldsymbol{p}^i_{pp_s}$ are the components of the s-th plane landmark $\boldsymbol{\Pi}_s$ of the i-th frame represented in the i-th frame.

8 Experiment

This section describes a series of experiments to analyze our proposed system. All the experiments are conducted on a PC equipped with an AMD Ryzen7 4800H CPU and 16G memory.

8.1 Dataset

The public dataset we used is UrbanLoco [16], an urban road dataset with a large number of pedestrians and vehicles. We only use the Hong Kong part of the dataset, the LiDAR used in this part is Velodyne HDL 32E, and the ground truth is provided by Novatel SPAN-CPT.

We also use a Robosense RS-LiDAR-16 to collect real-world indoor and outdoor datasets. We critically start and end the trajectory with the same point to evaluate the accuracy by calculating the difference between the estimated poses of the start and the end points.

8.2 Ablation Experiment

In order to test the different parts of our system, we evaluate the following variants of our system on the outdoor_1 sequence of our dataset and HK-Data20190316-2 sequence of UrbanLoco:

- MULO: Full set system.
- MULO-Undistort: Close motion compensation.
- MULO-Cylinder: The cylinder landmark is ignored.
- MULO-SuperWindow: Only the TinyWindow is used.
- MULO-BA: Only the front-end is used.

Table 1 is the result of the ablation experiment, which shows that our LiDAR BA method significantly improves the pose accuracy of the front-end. In the case of less plane information, only using plane landmarks maybe not work, and cylinder landmarks can avoid the effects of it. Because

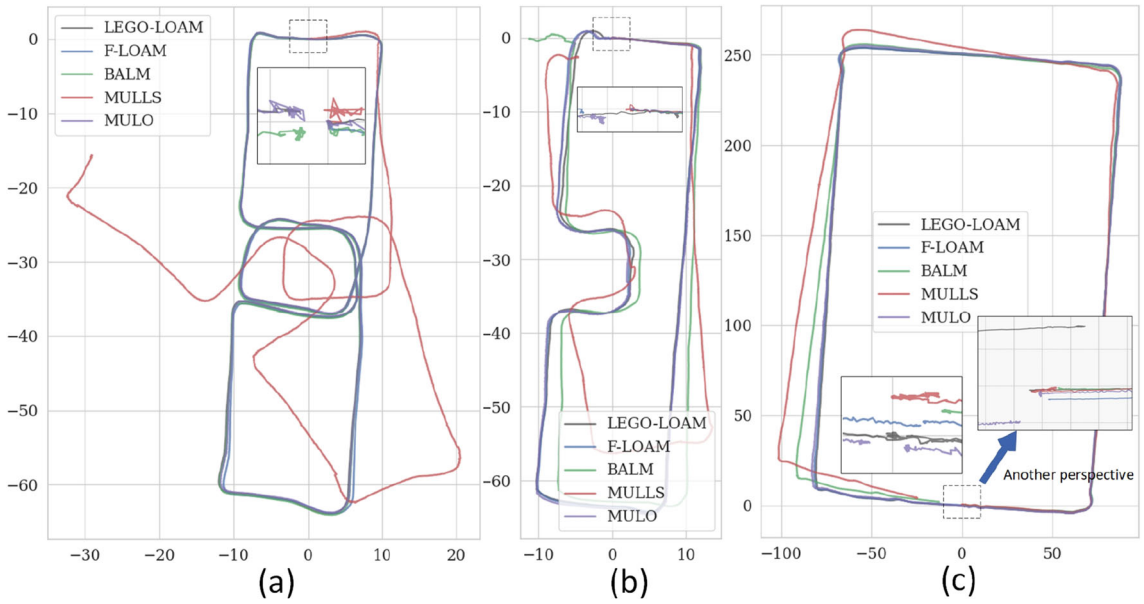

Fig. 6 Trajectories of our indoor and outdoor datasets. (a) indoor_1, (b) indoor_2, (3) outdoor_1

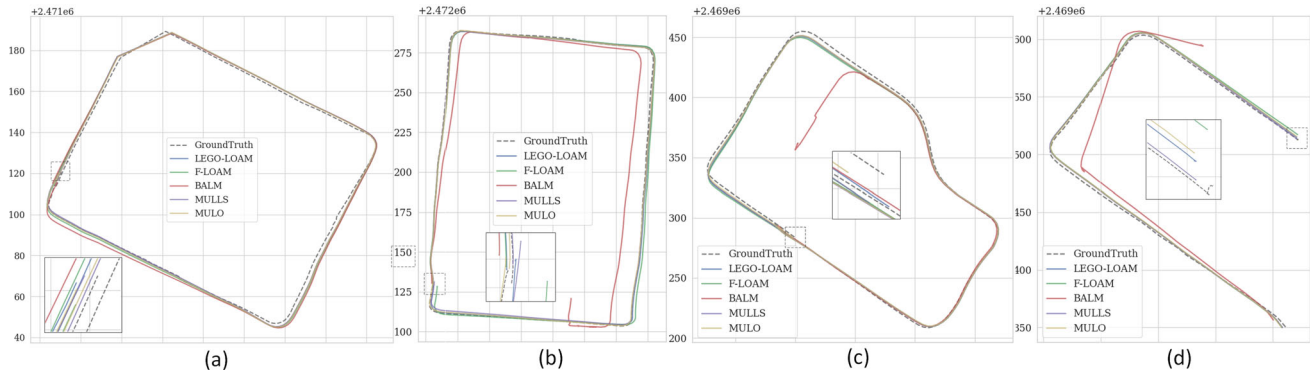

Fig. 7 Trajectories of the UrbanLoco dataset. (a) HK-Data20190426-2, (b) HK-Data20190316-2, (3) HK-Data20190117, (4) HK-Data20190426-1

LiDAR has a long measurement range, combining the SuperWindow with TinyWindow can make better use of the LiDAR information from a long distance. In addition, motion compensation also helps improve the system's accuracy.

8.3 Comparison with other Algorithms

We compare our system with LEGO-LOAM [3], F-LOAM [4], MULLS [5] and BALM [9]. The experimental results are shown in Tables 2 and 3. The trajectories are shown in Figs. 6 and 7. We modified the motion compensation module to a universal method for MULLS system and achieved higher accuracy on UrbanLoco Dataset. However, the results using the parameters provided by MULLS on our dataset are not ideal, which may be because the algorithm is not adapted for LiDAR with sparse point clouds. The results show that our proposed system achieves better accuracy than the other four state-of-the-art LiDAR SLAM systems except for HK-Data20190426-1 sequence, which is because dynamic removal is only used in the back-end, the bad prior pose affects the result of data association. A better front-end will have a better performance.

In addition, we can notice that the ICP-based methods such as LEGO-LOAM, F-LOAM, MULLS can work stably in various scenarios. This benefits from the fact that the ICP methods use a large number of points for alignment and does not require very accurate feature extraction, which makes it highly robust to plane missing scenes and dynamic scenes. However, this also leads to the ICP-based methods cannot further improve the trajectory accuracy. On the other hand, BALM, which performs BA optimization by accurately extracting plane and line landmarks, is highly vulnerable to dynamic and plane missing scenes due to its non-robust landmark extraction and data association methods. In contrast, our approach guarantees the robustness of landmark extraction and data association by a three-stage strategy. The use of cylinder landmarks further improves the system's accuracy and robustness. In addtion, our proposed new sliding window structure also gives the system higher accuracy than BALM in static structured scenes.

8.4 Runtime Analysis

Table 4 shows the runtime of MULO, which is tested on outdoor_1 sequence of our dataset with a Robosense RS-LiDAR-16 LiDAR. Our front-end achieves real-time, which takes 76ms for one scan. The back-end takes 143ms per round, but it is executed in a separate thread and only for keyframe. Therefore, our system can reach real-time on a moderate PC.

9 Conclusions

This paper proposes a novel lidar odometry MULO, which performs BA optimization using plane and cylinder landmarks. In addition, we present a new three-stage landmark extraction and data association strategy for plane and cylinder, which is efficient and robust. Furthermore, a new calculating stable parameterization method for cylinder landmark is proposed to model cylinder in optimization stage. A proposed sliding window that consists of a TinyWindow and a SuperWindow is used to conduct BA optimization with the two kinds of landmarks and scan poses. Finally, MULO is tested on UrbanLoco dataset and our dataset, and the results

Table 4 Time consumption for each module of our system. (ms)

Stage		Time consumption
Front-end	Feature points extraction	31
	Direct pose estimation	45
Back-end	Date association & New landmarks extraction	44
	Sliding window optimization	99

show that our system is competitive compared with the state-of-the-art LiDAR odometrys.

Author Contributions In this work, Wei Liu (corresponding author) coordinated the overall project and group activities. Jun Liu was responsible for the theoretical derivation and code writing, and Zhengnan He was responsible for the code writing and experimental design, and both of them jointly proposed the innovative points and wrote the theoretical derivation part of the manuscript. Zhao Xiaoyu was responsible for the experimental platform construction and experiments, and wrote the experimental part of the manuscript. Jun Hu and Shuai Cheng supplemented the experiments and wrote the other parts of the manuscript.

Funding Not applicable

Availability of Data and Materials Not applicable

Code Availibility Not applicable

Declarations

Conflict of Interest The authors have no relevant financial or non-financial interests to disclose

Ethics Approval No ethical approval is required by this research.

Consent to Participate Not applicable

Consent for Publication This paper does not require any consent for publication.

References

1. Besl, P.J., McKay, N.D.: Method for registration of 3-d shapes. In: Sensor Fusion IV: Control Paradigms and Data Structures, vol. 1611, pp. 586–606 (1992). Spie
2. Zhang, J., Singh, S.: Loam: lidar odometry and mapping in real-time. In: Robotics: Science and Systems, vol. 2, pp. 1–9 (2014). Berkeley, CA
3. Shan, T., Englot, B.: Lego-loam: lightweight and ground-optimized lidar odometry and mapping on variable terrain. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4758–4765 (2018). IEEE
4. Wang, H., Wang, C., Chen, C.-L., Xie, L.: F-loam: fast lidar odometry and mapping. In: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4390–4396 (2021). IEEE
5. Pan, Y., Xiao, P., He, Y., Shao, Z., Li, Z.: Mulls: versatile lidar slam via multi-metric linear least square. In: 2021 IEEE International Conference on Robotics and Automation (ICRA), pp. 11633–11640 (2021). IEEE
6. Rusu, R.B., Marton, Z.C., Blodow, N., Beetz, M.: Learning informative point classes for the acquisition of object model maps. In: 2008 10th International Conference on Control, Automation, Robotics and Vision, pp. 643–650 (2008). IEEE
7. Rusu, R.B., Blodow, N., Beetz, M.: Fast point feature histograms (fpfh) for 3d registration. In: 2009 IEEE International Conference on Robotics and Automation, pp. 3212–3217 (2009). IEEE
8. Rusu, R.B., Bradski, G., Thibaux, R., Hsu, J.: Fast 3d recognition and pose using the viewpoint feature histogram. In: 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2155–2162 (2010). IEEE
9. Liu, Z., Zhang, F.: Balm: bundle adjustment for lidar mapping. IEEE Robotics Automation Lett. **6**(2), 3184–3191 (2021)
10. Geneva, P., Eckenhoff, K., Yang, Y., Huang, G.: Lips: lidar-inertial 3d plane slam. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 123–130 (2018). IEEE
11. Ferrer, G.: Eigen-factors: plane estimation for multi-frame and time-continuous point cloud alignment. In: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 1278–1284 (2019). IEEE
12. Zhou, L., Koppel, D., Ju, H., Steinbruecker, F., Kaess, M.: An efficient planar bundle adjustment algorithm. In: 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 136–145 (2020). IEEE
13. Zhou, L., Koppel, D., Kaess, M.: Lidar slam with plane adjustment for indoor environment. IEEE Robotics Automation Lett. **6**(4), 7073–7080 (2021)
14. Huang, H., Sun, Y., Wu, J., Jiao, J., Hu, X., Zheng, L., Wang, L., Liu, M.: On bundle adjustment for multiview point cloud registration. IEEE Robotics Automation Lett. **6**(4), 8269–8276 (2021)
15. Zhou, L., Wang, S., Kaess, M.: π-lsam: lidar smoothing and mapping with planes. In: 2021 IEEE International Conference on Robotics and Automation (ICRA), pp. 5751–5757 (2021). IEEE
16. Wen, W., Zhou, Y., Zhang, G., Fahandezh-Saadi, S., Bai, X., Zhan, W., Tomizuka, M., Hsu, L.-T.: Urbanloco: a full sensor suite dataset for mapping and localization in urban scenes. In: 2020 IEEE International Conference on Robotics and Automation (ICRA), pp. 2310–2316 (2020). IEEE
17. Zuo, X., Yang, Y., Geneva, P., Lv, J., Liu, Y., Huang, G., Pollefeys, M.: Lic-fusion 2.0: Lidar-inertial-camera odometry with sliding-window plane-feature tracking. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 5112–5119 (2020). IEEE
18. Lee, W., Yang, Y., Huang, G.: Efficient multi-sensor aided inertial navigation with online calibration. In: 2021 IEEE International Conference on Robotics and Automation (ICRA), pp. 5706–5712 (2021). IEEE
19. Xu, W., Zhang, F.: Fast-lio: a fast, robust lidar-inertial odometry package by tightly-coupled iterated kalman filter. IEEE Robotics Automation Lett. **6**(2), 3317–3324 (2021)

Jun Liu received the B.S. degree in Automation from Shenyang Aerospace University, Shenyang, China, in 2013, and the M.S. degree in Biomedical Engineering from Northeastern University, Shenyang, China, in 2015. He currently is a Ph.D. student with the College of Information Science and Engineering, Northeastern University, Shenyang, China. His research interests include LiDAR SLAM, intelligent robot and computervision.

Zhengnan He received the B.E. degree from Taiyuan University of Technology, Taiyuan, China. He is currently working toward the M.A. Eng. degree with the College of Information Science and Engineering, Northeastern University, Shenyang, China. His research interests is LiDAR SLAM.

Xiaoyu Zhao received the B.S. degree in the College of Mechanical and Vehicle Engineering, Taiyuan University of Technology, Taiyuan, China and received the master's degree in the College of Information Science and Engineering, Northeastern University, Shenyang, China. His research interest is Multi-sensor fusion SLAM.

Jun Hu received the Ph.D. degrees in computer science and engineering from Northeastern University, Shenyang, China, in 2023. deputy director of the automatic driving business line, Neusoft Reachauto Corporation (Shenyang), Senior Engineer. His current research interests include assisted driving/automatic driving, image processing, pattern recognition, vision based perception of complex traffic scenes.

Shuai Cheng received the M.S. and Ph.D. degrees from the Changchun University of Science and Technology, Shenyang, China, in 2012 and in 2016. His current research interests include deep learning, object tracking, trajectory prediction in complex traffic scenes.

Wei Liu received the M.S. and Ph.D. degrees in control theory and control engineering from Northeastern University, Shenyang, China, in 2001 and in 2005, respectively. He is currently a professor-level Senior Engineer with the Research Academy, Northeastern University. He is also the Director of the Intelligent Vision Laboratory, Neusoft Corporation, Shenyang. His research interests include computer vision, image processing, and advanced driver assistance systems.