Wei WEI, Xiaorui ZHU, Yi WANG, 2022. Novel robust simultaneous localization and mapping for long-term autonomous robots. *Frontiers of Information Technology & Electronic Engineering*, 23(2):234-245.
https://doi.org/10.1631/FITEE.2000358

Novel robust simultaneous localization and mapping for long-term autonomous robots

Key words: Simultaneous localization and mapping (SLAM); Long-term; Robustness; Light detection and ranging (LiDaR); Visual inertial LiDaR navigation (VILN)

Wei WEI
E-mail: weirui9003@gmail.com
ORCID: https://orcid.org/0000-0002-8998-045X

Introduction for the SLAM system

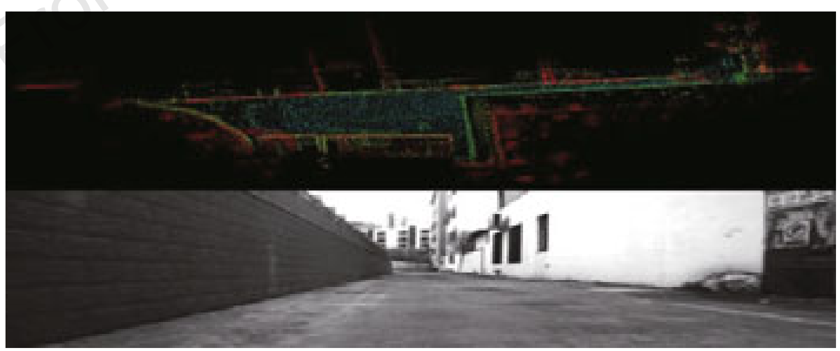
- Simultaneous localization and mapping (SLAM) is still a challenging problem for long-term autonomous mobile robots because the real world is full of highly dynamic, unstructured, and complex scenarios.

SLAM	Details
Google Cartographer (Hess et al., 2016)	Duplicated structural environments, such as long corridors or tunnels, might cause long-term robustness problems.
ORB-SLAM2 (Mur-Artal and Tardós, 2017)	It might not run for a long period in low-texture environments.
1-day learning, 1-year localization (Kim et al., 2019)	In new scenarios, the algorithm needs to learn the whole environment again.
Self-supervised deep pose corrections for robust visual odometry (Wagstaff et al., 2020)	It can hardly balance accuracy, efficiency, and long-term running compared with traditional SLAM systems.

Challenges

- For visual sensors, the rapid steering of the robot causes the image to be blurred, and some scenes cause over-exposure.
- A long corridor environment may cause a pure LiDaR system to fail.

Some crucial situations for pure visual SLAM systems



Long corridor environments

1) VILN SLAM framework

- VILN SLAM consists of three major systems, i.e., stereo visual inertial odometry (VIO), a visual-LiDaR mapping system, and a LiDaR enhanced visual loop closure system.

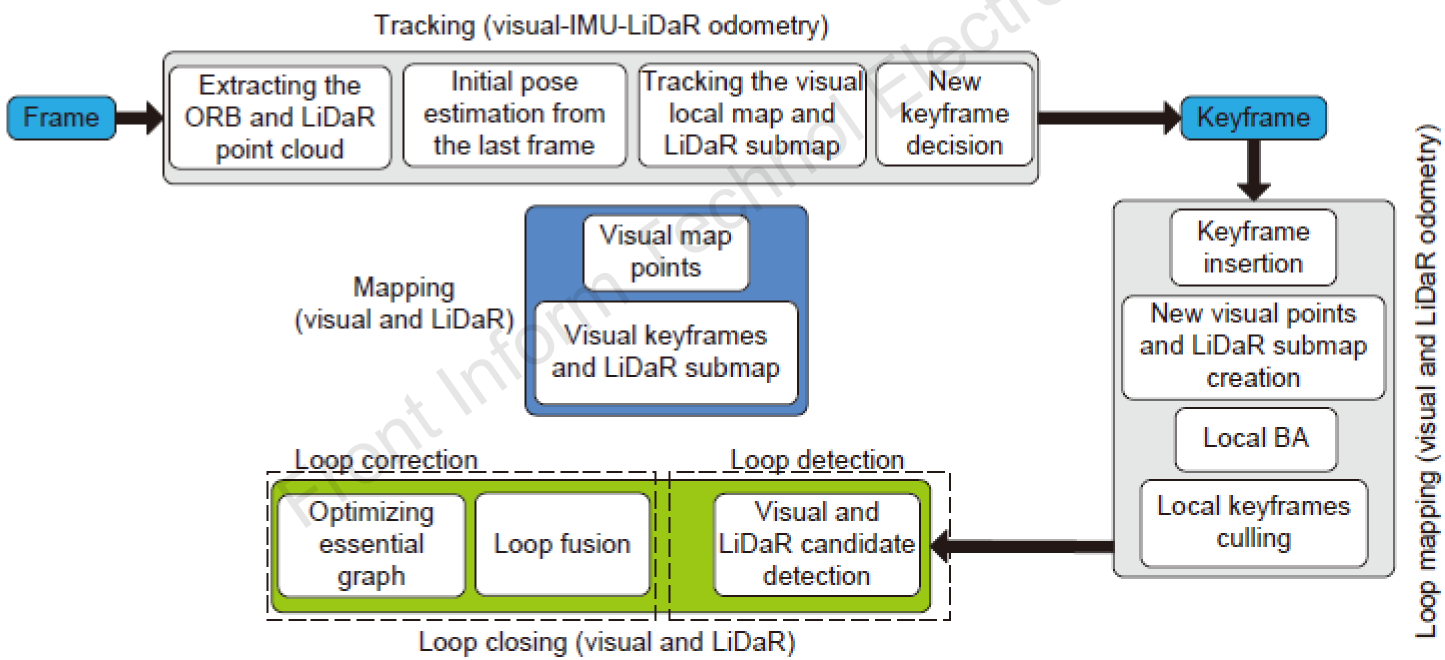
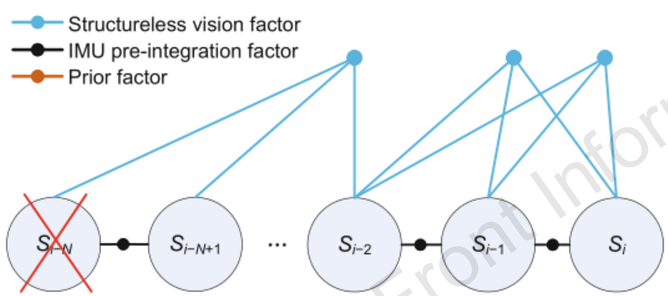


Fig. 1 Overview of the VILN SLAM framework

2) Stereo visual inertial odometry

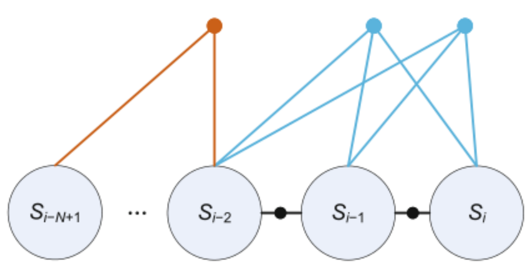
- Hybrid visual frontend: The system that combines the direct and feature-based methods becomes a hybrid of the two.
- Backend optimizer: Providing real-time locally consistent state estimate at a relatively high frequency is the goal of the backend optimizer, which will be served as the motion model for the LiDaR mapping algorithm.



(a)

$$S_W^* = \arg\min_{S_W^*} \Big(\|r_0\|^2_{\Sigma_0} + \sum_{i \in W} \|r^{\mathrm{I}}_{i(i+1)}\|^2_{\Sigma_{\mathrm{I}}} + \sum_{p} \|r^{\mathrm{V}}_{p}\|^2_{\Sigma_C} \Big)$$

Pose graph optimization



(b)

Pose-graph formulation in the visual inertial odometry

3) Visual LiDaR fused map and loop closure

- Introduce a visual-IMU-LiDaR fused mapping method, as well as a fast loop closing approach via the map.
- Use high-frequency IMU rate VIO poses as the motion prior to performing 3D LiDaR scan to map registration.

0	1	0	0	0	0	0	0	1	0
0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	1
0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0

1	0	1	2	2.236	1.414	1	1	0	1
1.414	1	1.414	2.236	2	1	0	1	1	1.414
2.236	2	2.236	2.828	2.236	1.414	1	0	1	2
3	3	3.162	3.606	2.828	2.236	1.414	1	1.414	1
2	2.236	2.828	3.162	3	2	1	0	1	0
1	1.414	2.236	2.236	2	2.236	1.414	1	1.414	1
0	1	2	1.414	1	1.414	2.236	2	2.236	2
1	1.414	2	1	0	1	2	3	3.162	3
1.414	1	1.414	1.414	1	1.414	2.236	3.162	4.123	4
1	0	1	2	2	2.236	2.828	3.606	4.472	5

Transforming a binary image to the Euclidean distance grid

3) Visual LiDaR fused map and loop closure (Cont'd)

- Scan matching algorithm and loop closing algorithm

Algorithm 1: Scan matching

```
best_score ← +∞
for (j_x, j_y, j_z, j_rx, j_ry, j_rz) ∈ W̄ do
    score ← Σ_{k=1}^{K} M_nearest(T_{ξ0+Δξ} h_k)
    Δξ = (r j_x, r j_y, r j_z, δθ j_rx, δθ j_ry, δθ j_rz)
    if score < best_score then
        match ← ξ0 + Δξ
        best_score ← score
return best_score and match
```

Where: score $\leftarrow \sum_{k=1}^{K} M_{\text{nearest}}(T_{\xi_0+\Delta\xi} h_k)$, $\Delta\xi = (rj_x, rj_y, rj_z, \delta_\theta j_{r_x}, \delta_\theta j_{r_y}, \delta_\theta j_{r_z})$, for $(j_x, j_y, j_z, j_{r_x}, j_{r_y}, j_{r_z}) \in \overline{\mathcal{W}}$, match $\leftarrow \xi_0 + \Delta\xi$.

Algorithm 2: Loop closing

```
if visual loop is detected (BoW) then
    if LiDaR loop is detected (Algorithm 1)
    then
        loop correction (graph optimization) by
        the LiDaR loop detection result
    else
        loop correction (graph optimization) by
        the visual loop detection result
else
    if LiDaR loop is detected (Algorithm 1)
    then
        loop correction (graph optimization) by
        the LiDaR loop detection result
    else
        loop correction fails
return success or failure
```

4) Experiments and discussion

- We have performed an extensive experimental validation of our VILN system in outdoor sequences from the KITTI dataset.

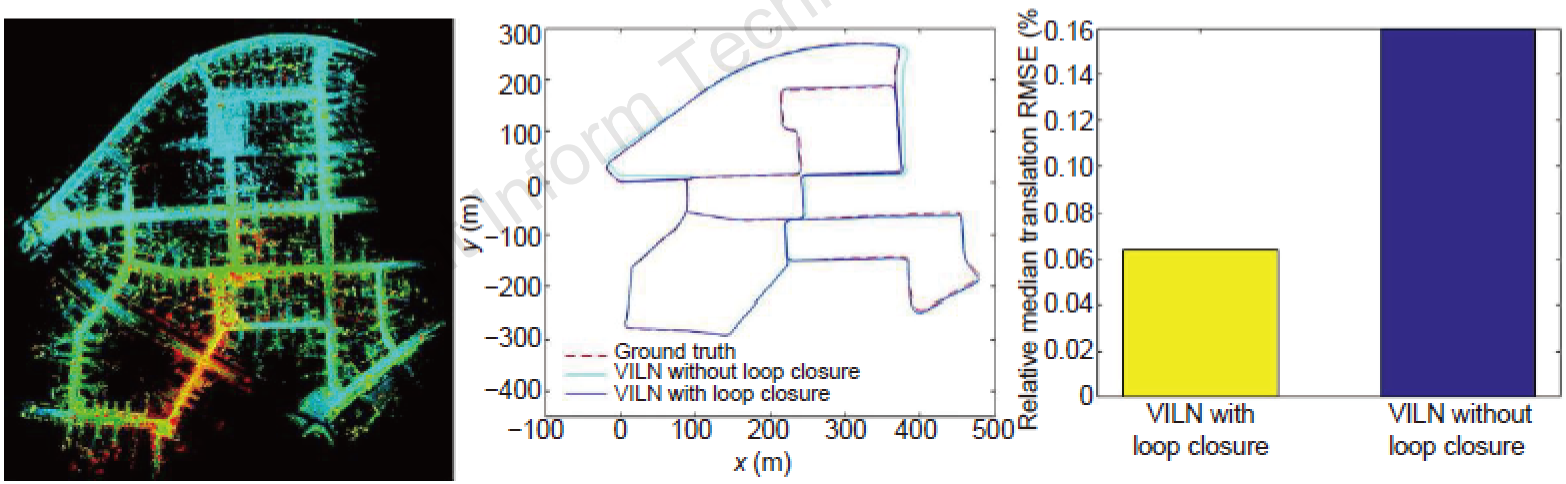
Table 1 Comparison of accuracy in the KITTI dataset

Sequence	Dimension (m×m)	Relative median translation RMSE (%)		Mileage (m)
		VILN	ORB-SLAM2	
KITTI 00	496×564	**0.064**	0.70	3744.90
KITTI 01	1157×1827	**0.19**	1.39	2461.75
KITTI 02	946×599	**0.091**	0.76	5251.44
KITTI 03	199×471	**0.23**	0.71	567.40
KITTI 04	394×0.5	**0.12**	0.48	393.43
KITTI 05	426×479	**0.11**	0.40	2224.17
KITTI 06	457×23	**0.17**	0.51	1236.30
KITTI 07	209×191	**0.088**	0.50	701.12
KITTI 08	391×808	**0.33**	1.05	3236.36
KITTI 09	568×465	1.67	**0.87**	1698.60
KITTI 10	177×671	**0.14**	0.60	924.93

Bold values represent better results

4) Experiments and discussion (Cont’d)

- The error of the system is typically below 0.33% no matter in small- or large-scale scenarios.
- In the future, we will focus on developing a new tightly coupled visual-IMU-LiDaR fusion navigation system to further improve the accuracy and long-term robustness under more complex and dynamic environments.



(a) (b) (c)

Fig. 6 Sequence 00 from the odometry benchmark of the KITTI dataset: (a) fusion maps; (b) ground truth and trajectories with and without loop closure; (c) relative RMSE comparison

References

[1] Hess W, Kohler D, Rapp H, et al., 2016. Real-time loop closure in 2D LIDAR SLAM. IEEE Int Conf on Robotics and Automation, p.1271-1278. https://doi.org/10.1109/ICRA.2016.7487258

[2] Mur-Artal R, Tardós JD, 2017. ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras. *IEEE Trans Robot,* 33(5):1255-1262. https://doi.org/10.1109/TRO.2017.2705103

[3] Kim G, Park B, Kim A, et al., 2019. 1-day learning, 1-year localization: long-term LiDAR localization using scan context image. *IEEE Robot Autom Lett* 4(2):1948-1955. https://doi.org/10.1109/LRA.2019.2897340

[4] Wagstaff B, Peretroukhin V, Kelly J, et al., 2020. Self-supervised deep pose corrections for robust visual odometry. IEEE Int Conf on Robotics and Automation, p.2331-2337. https://doi.org/10.1109/ICRA40945.2020.9197562

[5] Shao WZ, Vijayrangan S, Li C, et al., 2019. Stereo visual inertial LiDAR simultaneous localization and mapping. IEEE/RSJ Int Conf on Intelligent Robots and Systems, p.370-377. https://doi.org/10.1109/IROS40897.2019.8968012

[6] Geiger A, Lenz P, Stiller C, et al., 2013. Vision meets robotics: the KITTI dataset. *Int J Robot Res*, 32(11):1231-1237. https://doi.org/10.1177/0278364913491297

[7] Engel J, Koltun V, Cremers D, 2018. Direct sparse odometry. *IEEE Trans Patt Anal Mach Intell*, 40(3):611-625. https://doi.org/10.1109/TPAMI.2017.2658577

[8] Zhang J, Singh S, 2015. Visual-lidar odometry and mapping: low-drift, robust, and fast. IEEE Int Conf on Robotics and Automation, p.2174-2181. https://doi.org/10.1109/ICRA.2015.7139486

[9] Forster C, Carlone L, Dellaert F, et al., 2017. On-manifold preintegration for real-time visual-inertial odometry. *IEEE Trans Robot*, 33(1):1-21. https://doi.org/10.1109/TRO.2016.2597321

[10] Maurer CR, Qi RS, Raghavan V, et al., 2003. A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Trans Patt Anal Mach Intell*, 25(2):265-270. https://doi.org/10.1109/TPAMI.2003.1177156

魏伟，哈尔滨工业大学博士后，哈尔滨工业大学博士，美国Case Western Reserve University公派留学联合培养博士。主要研究方向包括无人驾驶技术、自主定位导航系统、自动控制等领域。已发表论文8篇，其中包括2篇*IEEE Trans Autom Contr*顶刊论文。

朱晓蕊，哈尔滨工业大学（深圳）教授、博士生导师。主要研究方向为移动机器人技术与系统。近年来作为负责人先后承担国家重点研发计划、国家科技支撑计划、国家自然基金重点项目等多项国家级前沿课题。曾任深圳大疆科技首席科学家、创始董事；孵化知名激光雷达厂商深圳速腾聚创并任首席科学家、联合创始人。

王毅，哈尔滨工业大学（深圳）教授、博士生导师。主要研究方向包括：
1. 新能源利用技术：三相大功率光伏并网发电系统控制策略；
2. 电机驱动控制技术：感应电机软起动控制系统，同步电机矢量控制系统；
3. 电动车控制技术：电池、电容充放电变换器、电机能量回馈、弱磁控制；
4. 开关电源技术：DC/DC变换器、AC/DC变换器。