ICPR 2020

25th International Conference on Pattern Recognition

AV-SLAM: Autonomous Vehicle SLAM with Gravity Direction Initialization

Participants

Kaan Yilmaz (1*), Baris Suslu (1*), Sohini Roychowdhury (2) and L. Srikar Muppirisetty (3)
1- Chalmers University of Technology, Department of Electrical and Computer Engineering, Göteborg, Sweden
2- Volvo Cars Technology USA, Mountain View, CA-940433 Volvo Car Corporation, SE-405 31 Göteborg, Sweden
3- Volvo Car Corporation, SE-405 31 Göteborg, Sweden

Project Description

Simultaneous localization and mapping (SLAM) algorithms map the environment of a vehicle by using on-board sensors (e.g., camera, Lidar, inertial measurement unit (IMU) etc.) and localize the vehicle within the extracted map. However, in practice, automotive sensors are susceptible to failures due to hardware wear and extreme weather conditions such as rain, snow, heat etc. To ensure adequate autonomous vehicle responses in such scenarios, a combination of multiple SLAM algorithms that has fail-safe modes allows the overall SLAM algorithm to function in the absence of one or several sensors. Therefore, we present a combination of three different localization algorithms in this work, where the different sensor modalities are: (i) IMU,(ii) camera and Lidar, and (iii) camera, Lidar and IMU.

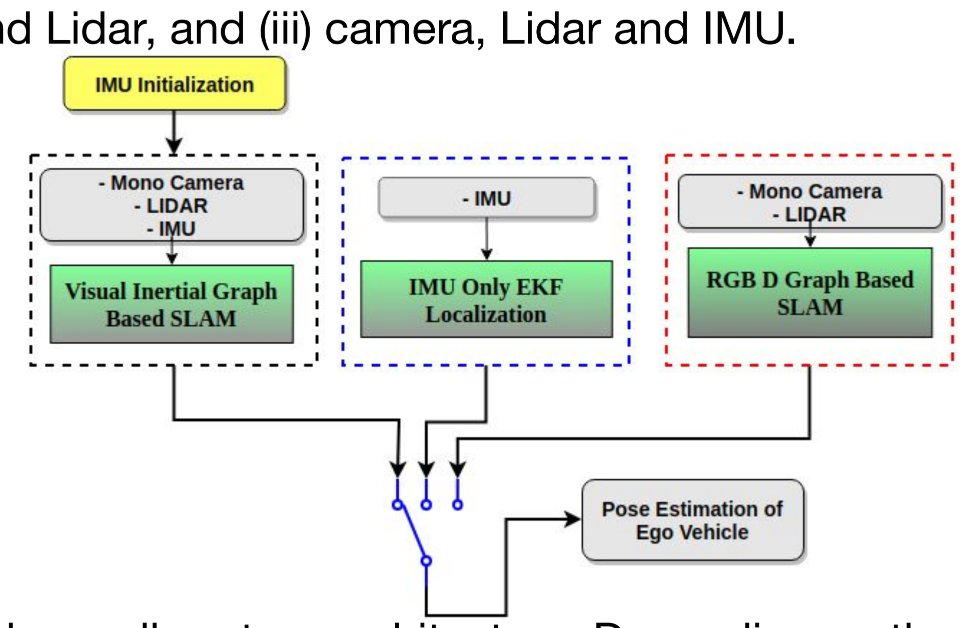

Fig 1: Proposed overall system architecture. Depending on the sensor availability, AV-SLAM can rely on different SLAM algorithms.

Whenever IMU is part of sensor configuration, it becomes essential to develop a robust initialization module that prevents divergent pose estimates made by the SLAM algorithm. Although IMU initialization has recently gained attention for accurate and computationally efficient pose estimation of unmanned aerial vehicle systems (UAVs) existing methods have not been shown to scale effectively from UAVs to AVs.

The purpose of this work is to present three localization algorithms that, in combination, can overcome unavoidable sensor failures. Also, a novel acceleration-based gravity direction initialization (AGI) method that is faster, more generalizable and more repeatable than the state-of-the art work in [1] for ego-vehicles operating under 38 km/h is presented. A novel SLAM framework suitable for the AV sensor requirements and fail-safe modalities is shown in Fig 1.

Theory and Method

The research has been splitted into three main parts as follows:

- **IMU-Only Localization**

In case where both camera and Lidar fail, localization can still be achieved with IMU-only EKF-localization. Therefore, Extended Kalman Filter (EKF) is applied to estimate the pose of the ego-vehicle where a fictitious measurement equation is given in the form of equality constraint. Here, we consider the lateral and the vertical velocities of the car (in IMU frame) to be zero. There are two pseudo measurements subjected to zero mean Gaussian noises. It should be noted that, in this algorithm, the initial IMU orientation is assumed to be known.

- **RGBD-SLAM**

RGBD-SLAM which is a Graph-SLAM approach, wherein a gray-scale monocular camera is used for sparse map extraction and a Lidar is used to estimate the scale of the extracted map. The architecture of the algorithm is depicted in Fig. 2.

- **Visual-Inertial SLAM (VI-SLAM)**

VI-SLAM is an extension of RGBD-SLAM, where IMU measurements are fused together with the pose estimated by RGBD-SLAM (see Fig. 2). The two main components of the integration step are IMU gravity direction initialization, and coupling of visual and inertial pose estimations.

- **Acceleration-based Gravity Direction Initialization Method (AGI)**

The proposed AGI method is based on estimating the gravitational acceleration by calculating the visually-estimated accelerations and subtracting them from raw accelerometer readings. This subtraction gives the acceleration due to gravity where roll and pitch angles become observable.

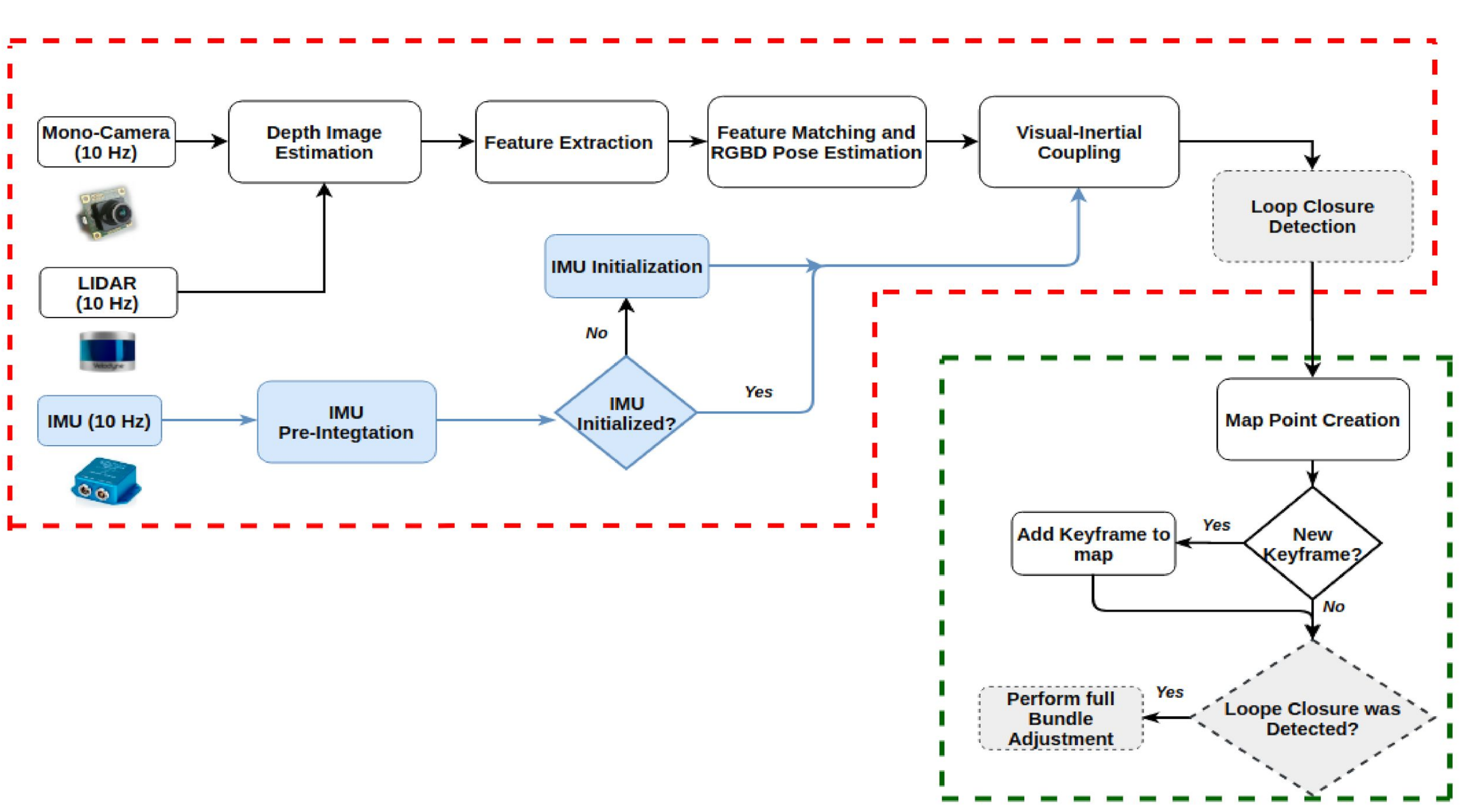

Fig 2: Overall architecture for RGBD-SLAM (white blocks) and VI-SLAM (white and blue blocks). Gray blocks represent loop closure support that can be integrated into the proposed architecture in future works

Results

The proposed localization algorithms are evaluated on raw KITTI odometry sequences [2]. Fig 3 shows the ground truth path and the estimated path by VI-SLAM for KITTI sequence 07. Resultant percentage translational relative pose errors (RPE) for all KITTI sequences are shown in Table 1. VI-SLAM performs best among the proposed methods and it outperforms ORB-SLAM2 and LOAM in select datasets. To evaluate the performance of AGI, the absolute initialization error of roll and pitch angles are calculated for each KITTI odometry sequence and they are compared with the gravity direction initialization method proposed in VINS-Mono [1] and a baseline method where the roll and pitch angles of the IMU is directly initialized to zero degrees. AGI outperforms other methods in datasets 7 8 9 and 10 and mostly results in less than 1 degree absolute initialization error.

Table 1: Comparison of percentage translational RPE of proposed algorithms

Seq No	% RPE_{trans}								
	IMU-Only (Proposed)	RGBD-SLAM (Proposed)	VI-SLAM			LOAM	ORB-SLAM2	CubeSLAM	PL-SLAM
			Baseline	AGI (Proposed)	VINS-Mono				
00	634	1.62	2.0	2.03	2.13	0.78	**0.7**	1.97	2.38
01	115	31.1	30.6	32.1	30.7	1.43	**1.39**	-	3.23
02	268	1.25	1.18	1.19	1.13	0.92	**0.76**	2.48	2.2
04	2.03	0.9	1.06	1.16	4.44	0.71	**0.55**	1.12	1.57
05	97.1	1.53	1.23	1.31	2.82	0.57	**0.37**	1.64	1.67
06	108	1.69	1.24	1.36	2.78	0.65	**0.43**	2.26	2.02
07	85.9	2.01	1.17	1.21	2.34	0.63	**0.45**	1.63	1.57
08	206	1.78	1.35	1.3	2.73	1.12	**1.06**	2.05	2.42
09	146	1.05	0.66	**0.65**	2.65	0.77	0.83	1.66	1.49
10	85.2	1.04	0.75	0.75	3.19	0.79	**0.55**	1.46	1.61

Table 2: Convergence plot of AGI. Green and red lines are the estimated and ground truth roll and pitch angles for gravity direction

Seq No	Roll - Abs Error (deg)			Pitch - Abs Error (deg)		
	Baseline	AGI (Proposed)	NS-Mono [6]	Baseline	AGI (Proposed)	NS-Mono [6]
00	2.41	0.8 ± 0.03	$\mathbf{0.57 \pm 0.1}$	1.27	2.38 ± 0.05	$\mathbf{1.06} \pm 0.02$
01	**0.73**	1.07 ± 0.95	9.6 ± 0.06	**3.01**	6.78 ± 0.45	4.44 ± 0.03
02	0.9	0.68 ± 0.01	$\mathbf{0.19 \pm 0.01}$	0.23	0.34 ± 0.05	$\mathbf{0.19} \pm 0.08$
04	**2.44**	3.22 ± 0.06	-	**0.66**	1.3 ± 0.08	-
05	**1.87**	3.85 ± 0.01	-	1.22	$\mathbf{0.09 \pm 0.02}$	-
06	**2.65**	5.3 ± 0.01	5.42 ± 0.02	**0.55**	1.38 ± 0.22	0.93 ± 0.01
07	1.29	$\mathbf{0.38 \pm 0.06}$	3.35 ± 0.2	**0.64**	0.93 ± 0.06	0.73 ± 0.01
08	2.95	$\mathbf{0.49 \pm 0.14}$	8.52 ± 0.3	1.37	$\mathbf{0.52 \pm 0.04}$	1.5 ± 0.02
09	1.92	1.46 ± 0.05	$\mathbf{0.94} \pm 0.05$	1.76	$\mathbf{0.1 \pm 0.01}$	1.83 ± 0.005
10	1.4	$\mathbf{0.82 \pm 0.03}$	-	3.22	$\mathbf{0.33 \pm 0.05}$	-

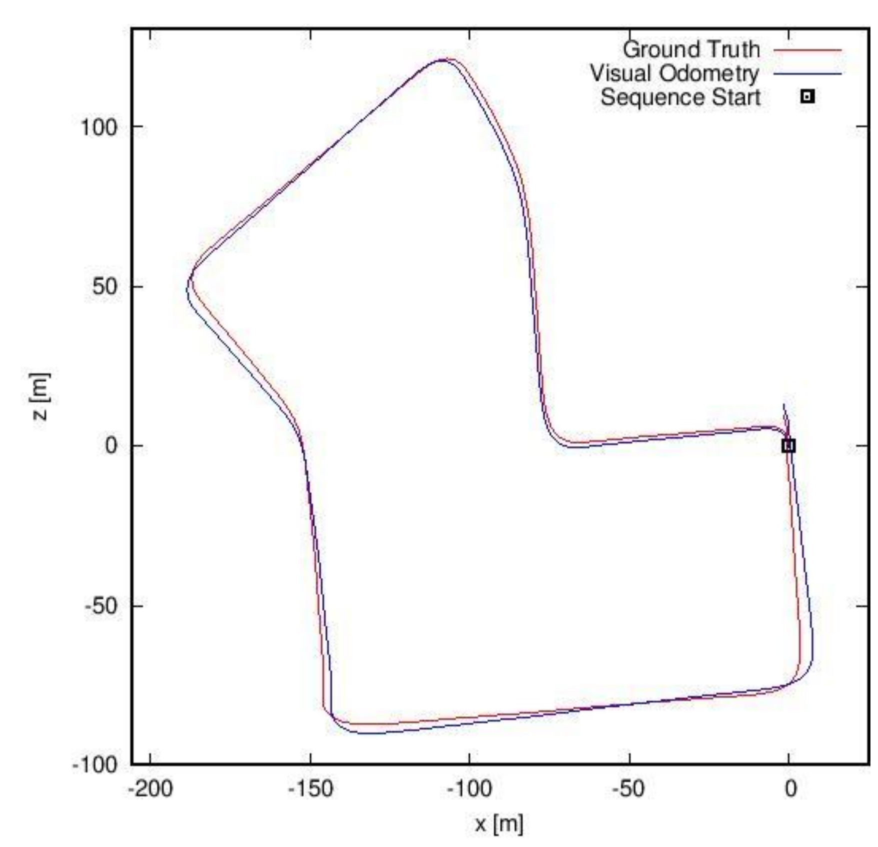

Fig 3: Ground truth path (KITTI Seq. 07) vs estimated path by VI-SLAM initialized with proposed AGI method

Conclusion

We present a novel SLAM framework that relies on IMU, camera and depth sensors and a robust gravity initialization method to enable fault tolerant performance in the event of sensor failures. From our analysis, we formulate three major conclusions:

1. The proposed framework ensures relative translational percentage error in pose of <2.03% using visual and combination with IMU sensors for low to medium speed ego-vehicle scenarios (speed <38 km/hr). SLAM algorithm with IMU-only sensor suffers from higher degrees of pose error with respect to the visual sensors owing to sensor biases.
2. The proposed SLAM framework with AGI method without loop closure modules outperforms state-of-the-art methods with loop closure modules on select data sequences. In datasets 09 and 10, the proposed VI-SLAM with AGI method results in RPE (pitch/roll initialization errors in degrees) of 0.65 (1.46/0.1) and 0.75 (0.82/0.33), respectively, which is an improvement over the method in LOAM with RPE of 0.77 and 0.79, respectively.
3. The proposed VI-SLAM with AGI provides faster and more reliable and repeatable initialization when compared to the existing method in.

References

[1] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," IEEE Transactions on Robotics, vol. 34, no. 4, pp. 1004–1020, 2018.

[2] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," The International Journal of Robotics Research, vol. 32, no. 11, pp. 1231–1237, 2013.

*Both authors have equal contribution