Seamless Outdoor and Indoor Mapping using a LiDAR-based Multi-Sensor System: Case Study on Kalman-Filter Fusion

ARASH JAVANMARD-GH[1], CHARLES TOTH[2] & DOROTA IWASZCZUK[1]

Abstract: Due to recent developments in the field of autonomous systems, pedestrian navigation, and Building Information Modeling, increasing interest in 3D indoor maps has been observed. One of the main challenges in indoor modeling is registration with outdoor models. A possible solution is seamless outdoor-indoor mapping using a mobile system operating well in both environments. In this paper, we introduce and compare some indoor and outdoor mapping methods for fusing outputs of a mobile multi-sensor system, such as IMU and LiDAR, where we enhance classical Extended Kalman Filter (EKF) based methods. In the first place, we will show the results of a loosely coupled error-based EKF for navigation, where we exploit the IMU measurements for prediction and the output of the iterative closest point (ICP) algorithm on the point cloud of the LiDARs for correction. Furthermore, we will also examine and compare the leverage of different ICP algorithms on the task of pose estimation.

1 Introduction

Accurate and georeferenced 3D maps of built areas have been gaining in importance over past years. Such concepts as smart cities rely on up-to-date models of the built-up world. Further-more, modern navigation solutions, including navigation of continuously growing number of autonomous systems and seamless pedestrian navigation, require accurate and georeferenced maps, including 3D indoor scenes. Indoor mapping is also of high interest for such applications as building maintenance and Building Information Modeling (BIM).

Georeferencing based on Global Navigation Satellite System (GNSS) can be successfully applied in outdoor. In indoor or in some outdoor environments the GNSS signal is however too weak for reliable georeferencing. Many approaches allowing the creation of indoor models using inertial measurement units (IMU) and/or point cloud matching can be found in the literature. Creating the connection between indoor and outdoor maps, however, remains challenging. Therefore, we investigate the feasibility of seamless outdoor-indoor mapping using a mobile multi-sensor mapping system. For this purpose, we compare and analyze the results of different indoor and out-door mapping approaches using the KITTI (GEIGER et al. 2012) dataset and one where a mobile platform consisting of three laser scanners (LiDARs), six RGB cameras, two GPS receivers and an IMU (IWASZCZUK et al. 2019) was used to collect geometric and photogrammetric information of an urban environment, including road networks, vegetation, city furniture, and building facades together with detailed structures, such as windows and doors as well as indoor scene with multiple floors and rooms. Due to the lack of synchronization at the time of publishing this paper we will use just the LiDAR point cloud from this dataset.

1 Technische Universität Darmstadt, Remote Sensing and Image Analysis, Franziska-Braun-Straße 7, D-64287 Darmstadt, E-Mail: iwaszczuk@geod.tu-darmstadt.de
2 The Ohio State University Columbus, Dept. of Civil, Environmental and Geodetic Engineering, OH, USA

2 Related Works

There have been many studies on laser or camera-based localization and mapping recently. Pose-Graph based approaches (GRISETTI et al. 2010; HESS et al. 2016; YE & LIU 2017) formulate the task of localization and mapping as a Maximum a Posteriori (MAP), where a front-end builds up the graph by connecting nodes depending on the measurements and a back-end subsequently tries to find the optimal estimation using non-linear optimization frameworks (CADEN et al. 2016). Nevertheless, Kalman-Filter based approaches are still of interest, especially an error-state based Extended Kalman Filter (ES-EKF) is due to its simplicity and the fact that its computation does not depend on growing covariance-matrix is very suitable for applications where many different sensory sources with different updates rate have to be fused to produce a more accurate trajectory estimation. Excellent work for a vision-aided inertial navigation system based on ES-EKF can be find in (MOURIKIS & ROUMELIOTIS 2007), by integrating the IMU measurement in the prediction step and a correction based on the camera triangulation. ZHEN et al. (2017) use an ES-EKF to fuse the odometry information of an IMU with Gaussian Particle Filter measurement updates. In this paper, we build up on their idea. But in contrast to their work, we will use different Iterative Closest Point (ICP) variations at the measurement step. ICP is by far one of the most popular methods for matching multiple point clouds. In the basic form, ICP is an optimization problem that is solved iteratively to find a hopefully good estimation of rotation and translation, which align the point clouds. Even though there exist many different variants, we will use two popular formulations, point-to-point (BESL & MCKAY 1992) and point-to-plane (YANG & MEDIONI 1992). Furthermore, we will use the open3d-library for point cloud processing (ZHOU et al. 2018).

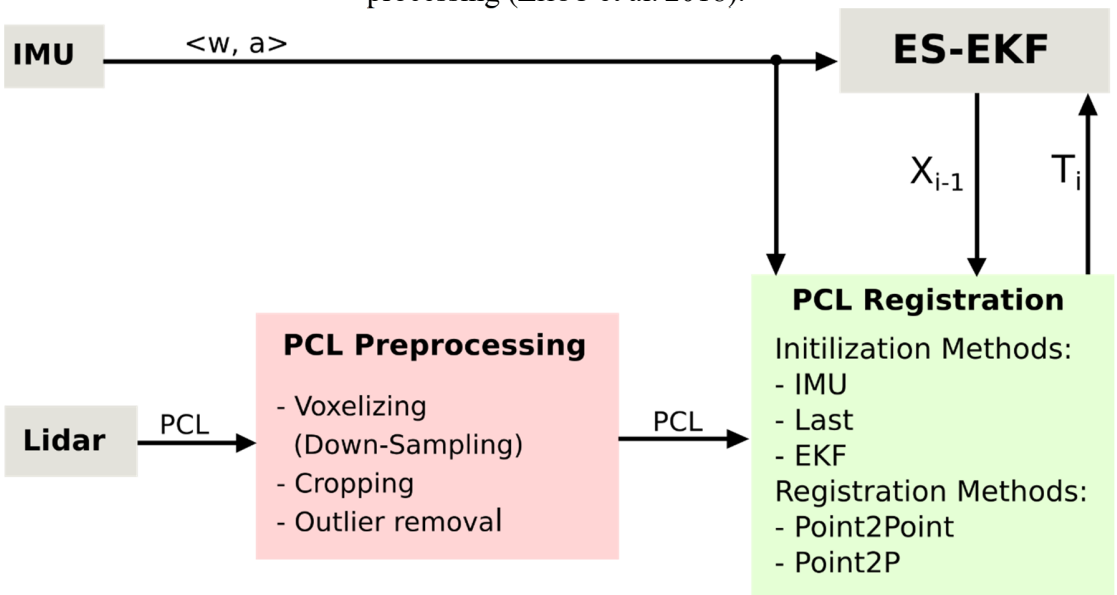

Fig. 1 gives an overview of the system, in section 3 we will give a description of the Point Cloud (PCL) processing and registration steps and in section 4 we will describe the procedure of ES-EKF for state estimation. Finally, in section 5 we present and discuss our results.

3 Point Cloud preprocessing and registration

In the preprocessing step the raw point cloud measurements are down sampled using a voxel grid filter based on its nearest neighbors (N) and a specific radius (r), which are both hyperparameters of the system and should be set a priori by the user. Since the noise of the lidar scanner increases with the distance of the points we also crop the point cloud by a spherical shape. The inner and outer radius of the shape are also hyper-parameters which should be chosen carefully by the user. At least a statistical outlier removal is applied to each point cloud. The adjusted output of the PCL-Processor is used by the PCL-Registration module. Like almost any other non-linear optimization-based approaches, also ICP needs an initial estimate of the transformation between two consecutive point clouds. We have implemented three approaches for this, the first initialization is done based on the assumption of constant velocity model, where we take the last estimate T_{i-1}. In the second case, we pre-integrate IMU measurements (FORSTER et al. 2015) between each LiDAR frame and calculated the initial transformation between two consecutive frames

$$T_{i-1,i}^{b} = T_{nb,i-1}^{-1} T_{nb,i} \tag{1},$$

where b is the body frame and n is the navigation frame.
Since IMU measurements suffer from noise by just integrating the measurements, our pose estimation will drift over time. To overcome this issue, we use the output of the ES-EKF to reset the last prediction of the IMU and start to integrate based on this corrected state.

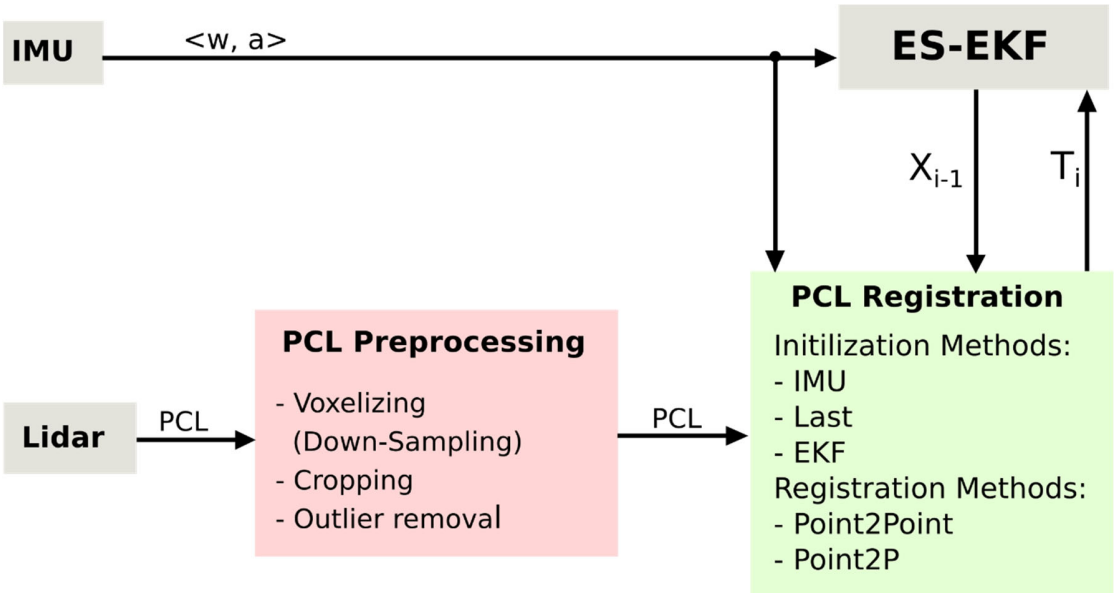

Fig. 1: System preview consisting of LiDAR preprocessing and registration, Extended Kalman Filter for state estimation and IMU measurements

4 IMU/Lidar fusion using ES-EKF

Dead reckoning is a widely used approach for pose estimation, wherein the goal is to find the optimal pose estimation by fusing different signals. One of the most widely used sensor combinations for pose estimation is IMU and Lidar. In the ES-EKF manner the linear acceleration and angular velocity measurements are integrated in the prediction step at the update rate of the IMU, which is typically 100-200 Hz and the output of a scan-matching algorithm which runs on much lower frequency 10Hz is fused for correction. Which is based on the LiDAR point clouds by calculating the transformation between each successive point cloud frames as it was elucidated in Sec. 3. In this work, we decided to use an Error-State EKF, due to its calculation simplicity and the fact the error states are close to zero, so that especially small-signal rotation approximation is linear and does not drift much from the nominal state (MADYASTHA et al. 2011)

$$R(\delta\theta) \approx I + [\delta\theta]_X \tag{2}.$$

Here is $[\delta\theta]_X$a skew-symmetric matrix of a small rotation $\delta\theta$.
Our formulation of the ES-EKF is based on the detailed investigative work in (SOLÀ 2017), but in contrast to that work we will not estimate the IMU bias in the state vector. Furthermore, we assume that our initial frame is a horizontal plane so we do not need to estimate the gravity vector at the initial frame. So that the nominal state dynamic vector $\dot{x}$ and the system error dynamics vector $\delta\dot{x}$ are represented as

$$\dot{x} = \begin{bmatrix} \dot{p} \\ \dot{v} \\ \dot{q} \end{bmatrix} = \begin{bmatrix} v \\ Ra_m + g \\ 0.5q \otimes w_m \end{bmatrix}, \quad \delta\dot{x} = \begin{bmatrix} \delta\dot{p} \\ \delta\dot{v} \\ \delta\dot{\theta} \end{bmatrix} = \begin{bmatrix} \delta\dot{v} \\ R[a_m]_x\delta\theta - a_n \\ w_m\delta\theta - w_n \end{bmatrix} \tag{3},$$

where $\dot{p}, \dot{v} \in R^3$denote velocity and acceleration vectors and $q \in H$is the quaternion representation of orientation. As mentioned before constant gravity vector g =$[0,0,9.81]^T$.Linear acceleration $a_m$and angular velocity $w_n$are considered as the control inputs. Furthermore $a_n, w_n$are accordingly system noise, which are assumed to be zero mean Gaussian random variables, e.g.,

$$a_n \sim N(0, Q_a), w_n \sim N(0, Q_w) \tag{4}.$$

The above in continuous time formulated differential equations should be discretized, since our sensor measurements are available just at some discrete time intervals δt. This can be done by any integration methods, like k^{th}-Order Runge-Kutta. We waive to formulate the whole integration step and refer to (SOLÀ 2017) for more information. In this work, we will use the Euler integration method, which is 1th-Order Runge-Kutta.

4.1 Prediction step

In the prediction, the nominal state is updated using the IMU measurements as control input as follow:

$$\check{x} = \begin{bmatrix} p_t \\ v_t \\ q_t \end{bmatrix} = \begin{bmatrix} p_{t-1} + v\Delta t + 0.5\big(R_{t-1}a_{m,t} + g\big)\Delta t^2 \\ v_{t-1} + (R_{t-1}a_m + g)\Delta t \\ q_{t-1} \otimes q(w_m \Delta t) \end{bmatrix} (5),$$

$$\begin{bmatrix} \delta p_t \\ \delta v_t \\ \delta \theta_t \end{bmatrix} = \begin{bmatrix} \delta p_{t-1} + \delta v \Delta t \\ \delta v - R[a_m]_x \delta\theta \Delta t + \hat{a_t} \\ R(w_m \Delta t)^T \delta\theta + \hat{w_t} \end{bmatrix} (6).$$

Error state itself can be estimated at the same time as measurement is observed, but nevertheless the error covariance P of the error state should be propagated at the same rate as IMU,

$$\check{P}_t = F_{t-1} P_{t-1} F_{t-1}^T + L_{t-1} Q_{t-1} L_{t-1}^T \quad (7).$$

With F and L being the transition matrix and the derivative of the discretized error-state dynamics w.r.t. the noise variables

$$F_t = \begin{bmatrix} I & I\Delta t & 0 \\ 0 & I & -R[a_m]_x \Delta t \\ 0 & 0 & R^T(w_m \Delta t) \end{bmatrix} Q_t = \begin{bmatrix} Q_a & 0 \\ 0 & Q_w \end{bmatrix} L_t = \begin{bmatrix} 0 & 0 \\ I & 0 \\ 0 & I \end{bmatrix} \quad (8).$$

We assume that the covariances Q_a, Q_w of the IMU are isotropic and can be read form the IMU datasheet.

$$Q_a = \sigma_a^2 \Delta t^2 I \; [m^2/s^2], Q_w = \sigma_w^2 \Delta t^2 I \; [rad^2] \quad (9).$$

The units are the results of integrating the covariances of zero mean Gaussian process variables (MAYBECK & SIOURIS 1980).

4.2 Correction step

Using the pose estimation of the PCL-Registration, the measurement model is formulated as

$$y_t = Hx_t + v_t, H = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \end{bmatrix} \in R^{6x9}, v \sim N(0, V_{lidar}) \quad (10).$$

Now by calculating all necessary information, the Kalman Gain and the error made during prediction step are

$$\begin{aligned} K_t &= P_t H_t^T (H_t P_t H_t^T + V)^{-1} \\ \delta\hat{x}_t &= K_t (y_t - \check{x}_t) \end{aligned} \quad (11),$$

In the Closed-Loop correction Kalman Filter (GROVES 2008) the errors estimated by the filter are fed back to the current nominal state and at the last step the error covariance is also corrected as follows

$$\begin{matrix} \hat{p} = \check{p} + \delta p_k \\ \hat{v} = \check{v} + \delta v_k \\ \hat{q} = q(\delta\theta) \otimes \check{q} \end{matrix} \quad , \hat{P} = (1 - K_t H_t)\check{P} \tag{12}.$$

5 Evaluation and Discussion

In this paper, we present preliminary results of our investigation by evaluating different combinations of the mentioned approaches in indoor and outdoor environments. We investigate, the quality of point cloud registration based on the of pure IMU mechanization, also we show how IMU can aid the ICP registration, especially when, due to a large transformation ICP does not converge to a good solution. Since IMUs are affected by noise, we examine the effect of Kalman Filter ICP registration. At least we compare two different registration methods, namely *poitn2point* and *point2plane* for outdoor to indoor transition.
Table 1 shows the results of the proposed approaches on the KITTI dataset, where a vehicle was driving for about 90 sec.
In Fig. 2 two different maps made by the proposed approaches are shown, on the right map a strong misalignment, due to drift around the roll axis can be seen. In contrast, however, in the left image, a map and a corresponding trajectory were built.

Tab. 1: Results of different approaches performed on ~90 [sec] of KITTI dataset.

Initialization	**Regist. Method**	**Position Error**	**Orientation Error**
Imu	Imu	31%	12%
Imu	Point2Plane	10%	9%
Imu	Point2Point	19%	19%
Imu + ES-KF	Point2Point	**3.5%**	3%
Last	Point2Plane	4.5%	8%
Last	Point2Point	4.5%	**2.8%**

We also test here described methodology on our outdoor-indoor dataset captured by the backpack mobile-mapping system (IWASZCZUK et al. 2019). The Velodyne VLP-16, which was used in this mapping system, consists of 16 lines, according to this, the point clouds collected at each revolution are sparse, unlike the LiDAR systems with many more lines, like one used in KITTI dataset. This sparsity makes an appropriate registration much harder, specially *point2plane* registration produces more outliers and bad registrations, since the radius for estimating the normals at each point must be large enough to make a plane estimation accurate enough. But on the other side, large radius may result in bad normal estimations since the points may not belong to the same surface anymore. As can be seen in Fig. 3 and Fig. 4, *poitn2point*-method produces fewer misalignments compared to *poitn2plane* (red area). Nevertheless, both approaches fail to keep a good estimate of the height; especially in an indoor area the floors are collapsed after some iterations.

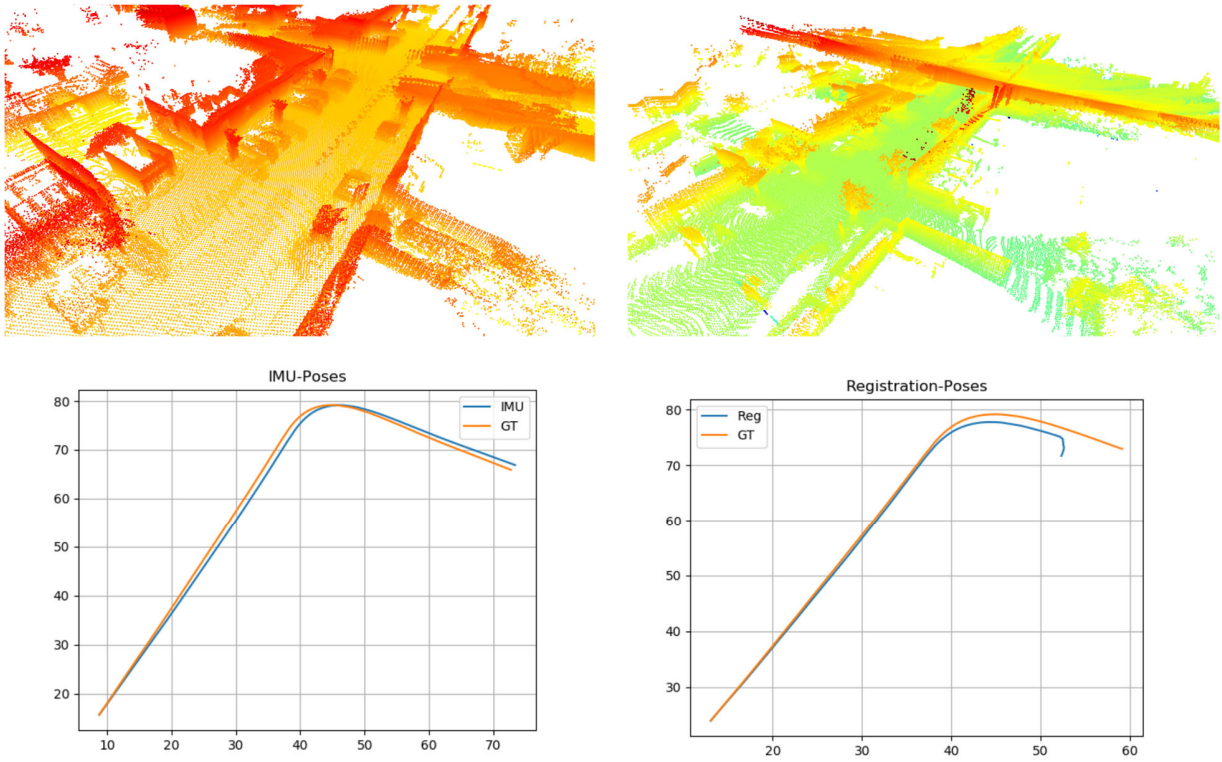

Fig. 2: Top: A 3D map of the scene, right) good registration results by IMU+EKF, left) Registration by pure IMU, where a drift about the roll axis caused a bad registration result. Bottom: Comparison of the estimated trajectory against ground truth

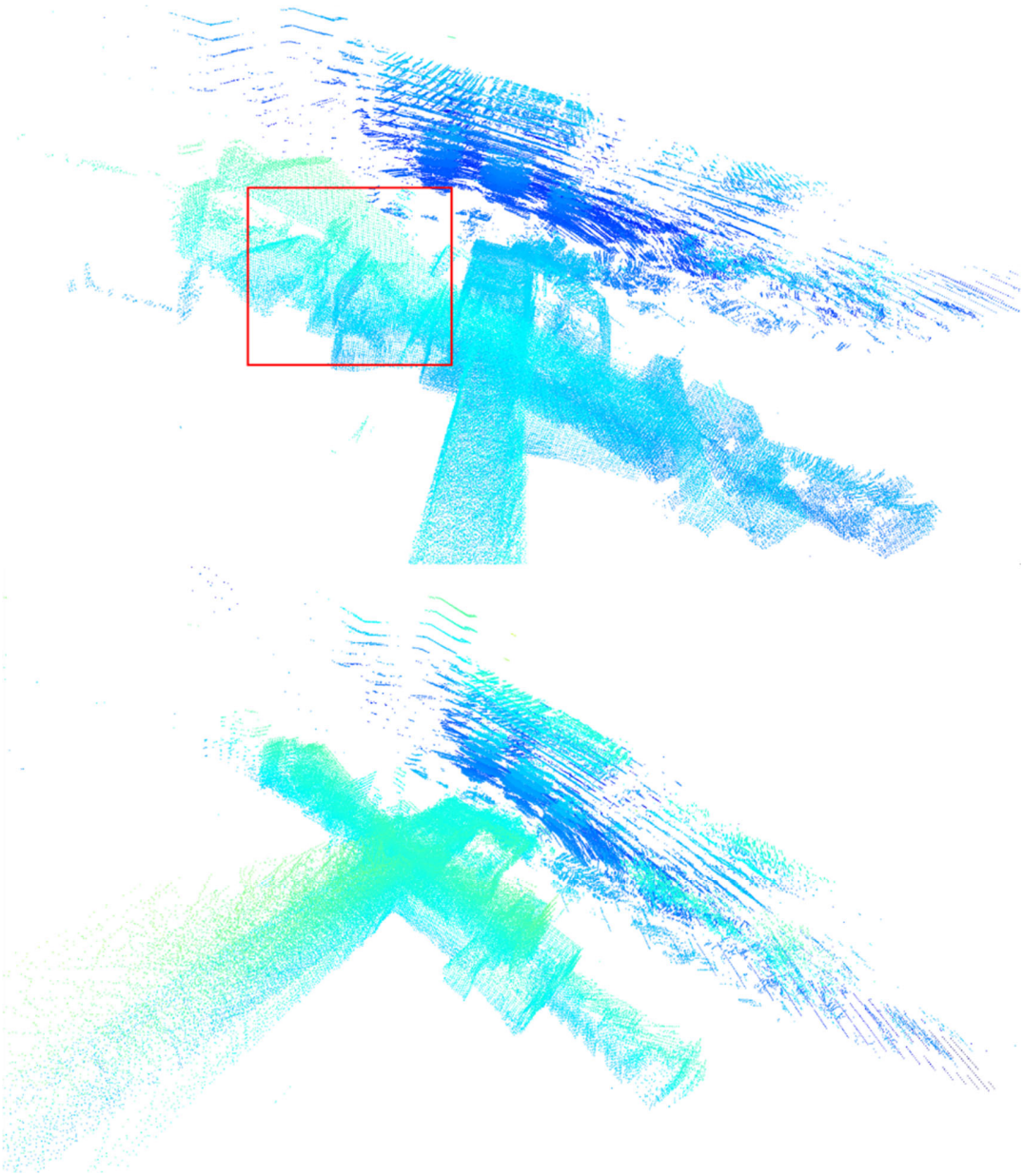

Fig. 3: 3D-Map created by *point2plane* registration (top), 3D-Map created by *point2point* registration (bottom)

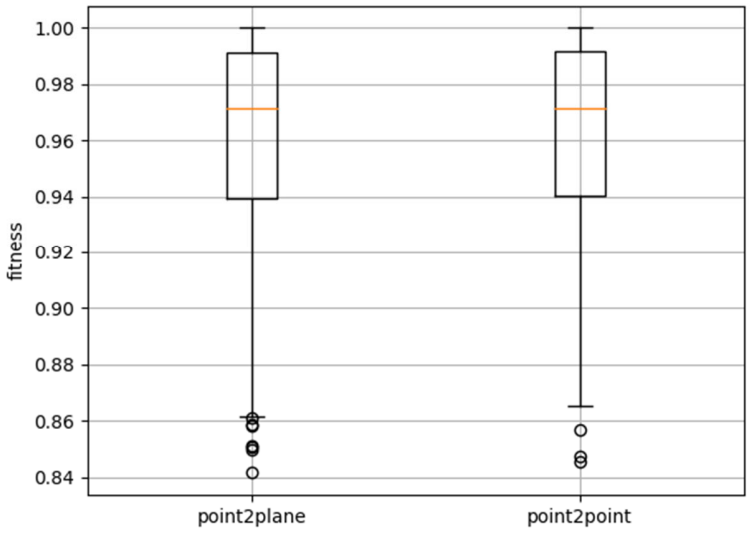

Fig. 4: Comparison between *point2point* and *point2plane* fitness values

6 Future Work

In this study, we have removed the estimation of the IMU bias from ES-EKF formulated in this paper, and assumed it is given and constant. This assumption is rather rarely true, unless we have a highly accurate navigation grade sensor. We expect that moving the IMU bias into the dynamics of the error state, provided that they are observable, would result in more accurate state estimation. This issue will be the subject of an in-depth future investigation. Moreover, one drawback of using EKF in both, loosely coupled or tightly coupled modes, is that due to the linearization and the presence of potentially strong non-linear effects, the map becomes more inconsistent over time. To overcome this issue, we plan to explore the possibilities of combining Pose Graph-based approaches; especially for building a more accurate map during the outdoor-indoor transition. Another issue we have observed during this work is the weakness of current ICP based methods when registering sparse Point Clouds, like those generated by Velodyne VLP-16. Therefore, we will also investigate the benefits of using more sophisticated feature-based registration methods. For example, we are going to perform an unsupervised classification based on the geometrical features extracted from point clouds. Afterward, feature association followed by registration is performed. We hope to get more satisfactory registration results than pure ICP. We also plan to investigate deep learning methods either as a part of the whole system, e.g. for registration, or in an end-to-end network for the whole task of odometry estimation.

7 References

BESL P. J. & MCKAY, N. D., 1992: A Method for Registration of 3-D Shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence **14**(2), 239-256.

CADEN, C., CARLONE, L., CARRILLO, H., LATIF, Y., SCARAMUZZA, D., NEIRA, J., REID, I. & LEONARD, J.J., 2016: Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. IEEE Transactions on robotics **32**(6), 1309-1332.

FORSTER, C., CARLONE, L., DELLAERT, F. & SCARAMUZZA, D., 2015: IMU preintegration on manifold for efficient visual-inertial maximum-a-posteriori estimation. In Robotics: Science and Systems, http://www.roboticsproceedings.org/rss11/p06.pdf.

GEIGER, A., LENZ, P. & URTASUN, R., 2012: Are we ready for autonomous driving? the KITTI vision benchmark suite. IEEE Conference on Computer Vision and Pattern Recognition, 3354-3361.

GRISETTI, G., KUMMERLE, R., STACHNISS, C. & BURGARD, W., 2010: A tutorial on graph-based SLAM. IEEE Intelligent Transportation Systems Magazine **2**(4), 31-43.

GROVES, P. D., 2008: Principles of GNSS Inertial and Multi-Sensor Integrated Navigation Systems. In GNSS Technology and Applications. 2008.

HESS, W., KOHLER, D., RAPP, H. & ANDOR, D., 2016: Real-time loop closure in 2D LIDAR SLAM. IEEE International Conference on Robotics and Automation (ICRA), 1271-1278.

IWASZCZUK, D., KOPPANYI, Z., PFRANG, J., TOTH, C., 2019: Evaluation of a Mobile Multi-Sensor System for Seamless Outdoor and Indoor Mapping. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. **42**(1/W2), 31-35.

MADYASTHA, V. K., RAVINDRAY, V. C., MALLIKARJUNAN, S. & GOYAL, A., 2011: Extended Kalman filter vs. error state Kalman filter for aircraft attitude estimation. AIAA Guidance, Navigation, and Control Conference, 6615.

MAYBECK, P. S. & SIOURIS, G. M., 1980: Stochastic Models, Estimation, and Control. Volume I, IEEE Transactions on Systems, Man and Cybernetics.

MOURIKIS, A. I. & ROUMELIOTIS, S. I., 2007: A multi-state constraint Kalman filter for vision-aided inertial navigation. IEEE International Conference on Robotics and Automation, 3565-3572.

SOLÀ, J., 2017: Quaternion kinematics for the error-state Kalman filter. arXiv:1711.02508.

YANG C. & MEDIONI, G., 1992: Object modelling by registration of multiple range images. Image Vision Comput. **10**(3), 145-155

YE, H., & LIU, M., 2017: LiDAR and Inertial Fusion for Pose Estimation by Non-linear Optimization. arXiv preprint arXiv:1710.07104.

ZHEN, W., ZENG, S. & SOBERER, S., 2017: Robust localization and localizability estimation with a rotating laser scanner. IEEE International Conference on Robotics and Automation (ICRA), 6240-6245.

ZHOU, Q.-Y., PARK, J. & KOLTUN, V., 2018: Open3D: A Modern Library for 3D Data Processing. arXiv preprint arXiv:1801.09847.