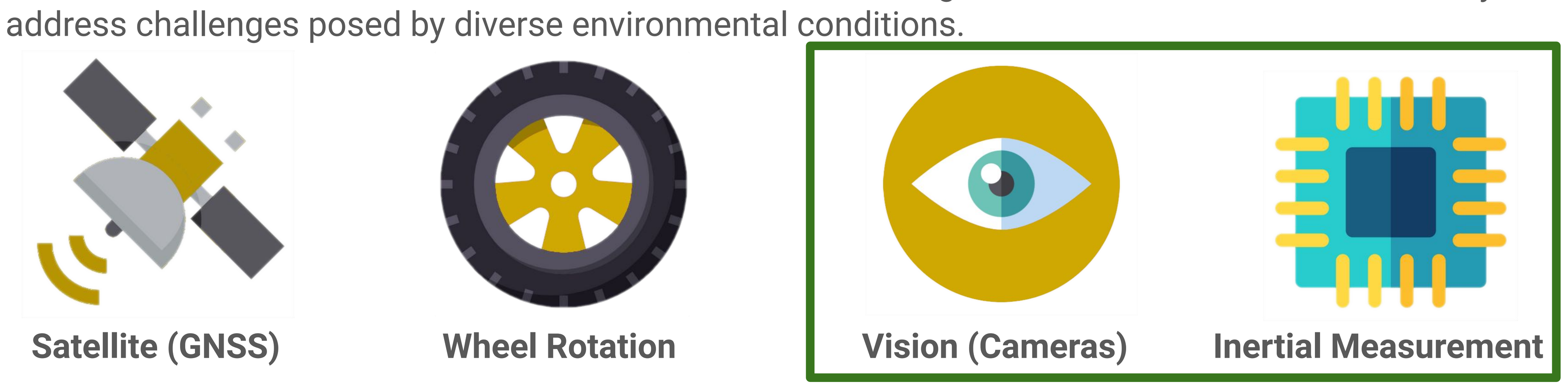
Georgia Tech Institute for Robotics and Intelligent Machines

Visual Inertial Odometry based State Estimation for Autorally

Advised by: Panagiotis Tsiotras; Poster by: Bhushan Pawaskar
(Dynamics and Control Systems Laboratory)

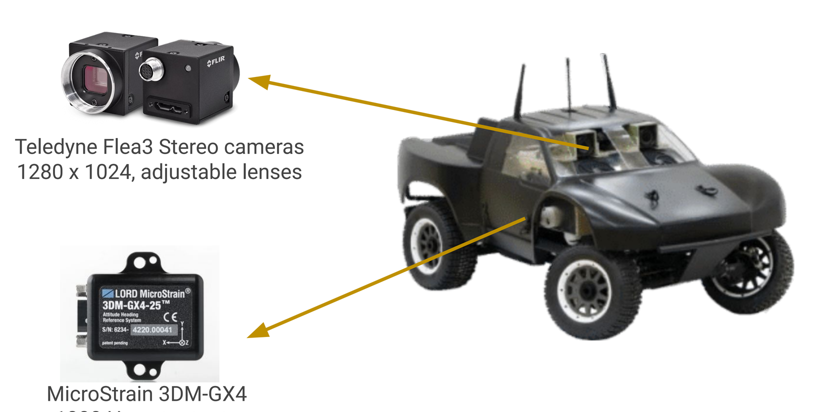
ABSTRACT

This project focuses on the task of **localization for Autorally** testbed environment **using visual camera data**. Traditional methods, like Global Navigation Satellite Systems (GNSS) have limitations, such as signal interference whereas IMU based estimation methods suffer from lots of noise in data collected at source. The project proposes a Visual Inertial Odometry (VIO) System that **combines the information from feature rich RGB camera data with IMU data**, aiming to enhance localization accuracy and address challenges posed by diverse environmental conditions.

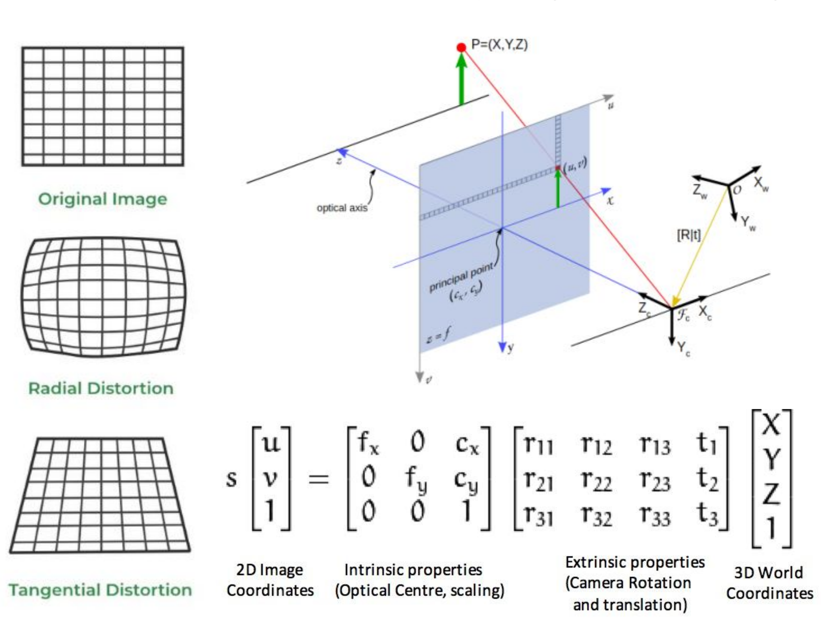


METHODOLOGY

Models to estimate state based on hardware readings are based on ideal assumptions. Simulated hardware models are ideally perfect. Real life hardware on the other hand has lots of imperfections. **Hardware requires calibration!**



Cameras need to account for projection and distortion errors (Camera intrinsic parameters). Pinhole camera model with radial and tangential distortion is chosen for its simplicity and accuracy.



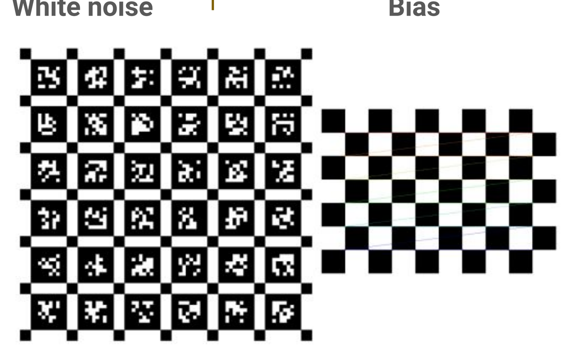
$$s\begin{bmatrix}u\\v\\1\end{bmatrix}=\begin{bmatrix}f_x&0&c_x\\0&f_y&c_y\\0&0&1\end{bmatrix}\begin{bmatrix}r_{11}&r_{12}&r_{13}&t_1\\r_{21}&r_{22}&r_{23}&t_2\\r_{31}&r_{32}&r_{33}&t_3\end{bmatrix}\begin{bmatrix}X\\Y\\Z\\1\end{bmatrix}$$

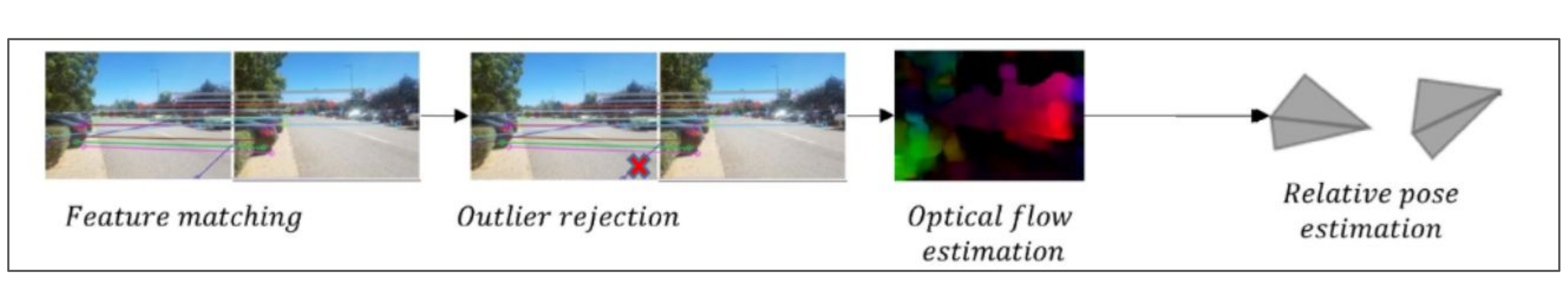
Inertial Measurement Units (IMU) parameters need to account for the noise in the IMU. This is usually given as a constant **bias and additive white noise** both for **gyroscope** as well as **accelerometer**.

White noise	Bias
$n_d[k]=\sigma_{g_d}w[k]$	$b_d[k]=b_d[k-1]+\sigma_{bgd}w[k]$
$w[k]\sim\mathcal{N}(0,1)$	$w[k]\sim\mathcal{N}(0,1)$
$\sigma_{g_d}=\sigma_g\frac{1}{\sqrt{\Delta t}}$	$\sigma_{bgd}=\sigma_{b_g}\sqrt{\Delta t}$

Both **Camera and IMU** need to work in tandem, so we also need to estimate their **relative transformations** with each other.

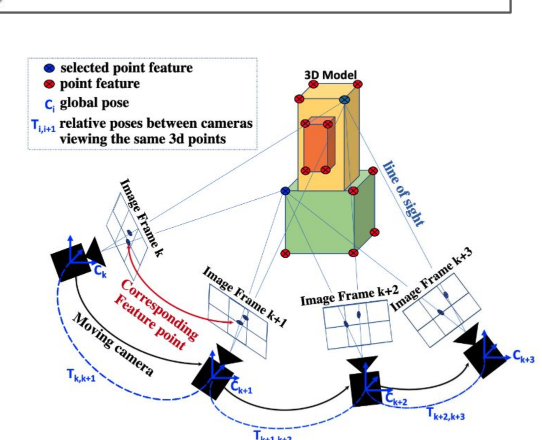
This is achieved using a **calibration sheet** as shown and exciting all axes of the IMU.

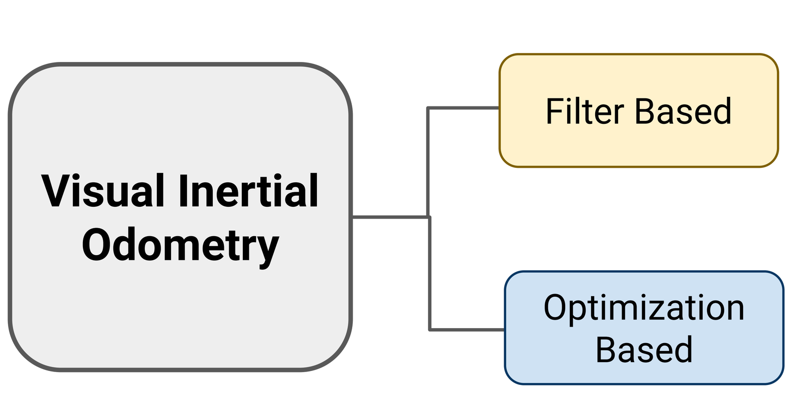




Visual Odometry systems work by **finding common feature points** between two poses. The algorithm solves an **optimization problem** that **minimizes reprojection error between poses** to find the most probable transformation between them.

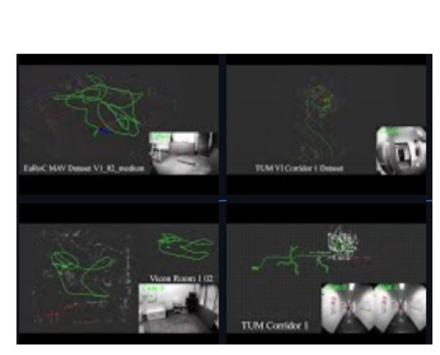
$$T_t^k=argmin_{T_t^k}\,\Sigma_i\|P_t^i-P_{t-1}^i\|^2$$





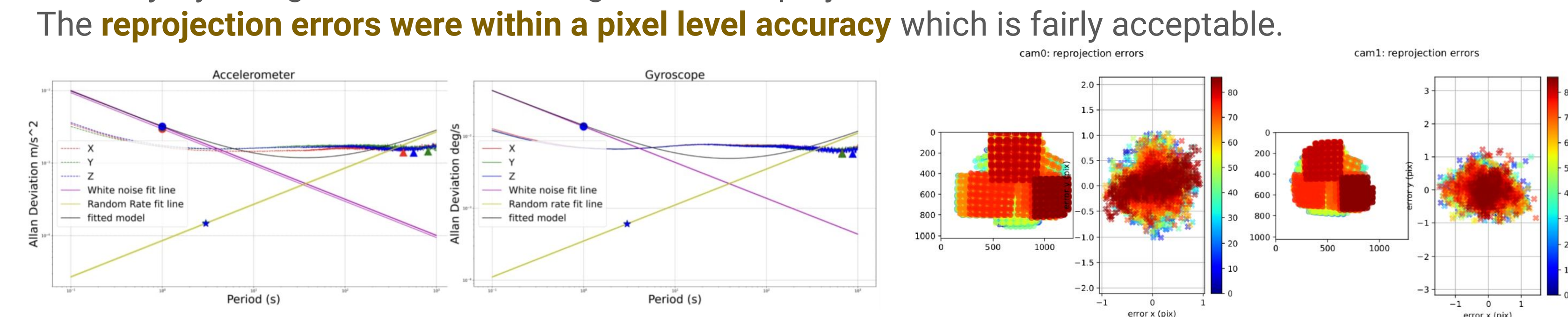
Step by Step update of inertial and visual measurements alternatively **using an Extended Kalman Filter** with previous estimate

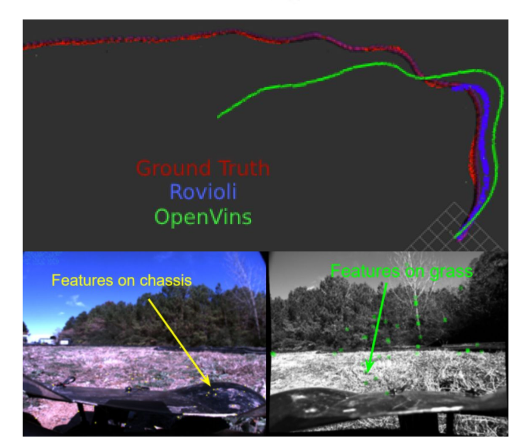
Solve **nonlinear optimization problem** over both inertial and visual measurements at once - **minimize relative squared error of estimate**



RESULTS

The IMU data was recorded keeping it stable for two to three hours. This got rid of the initial jerks and then only the IMU noise remained. These **allan deviations were plotted on a logarithmic time scale** for both accelerometer and gyroscope. After fitting a curve to these noise results, **white noise and bias sigmas were obtained** using the Allan Variance method.
Similarly by using the calibration target, camera projection and distortion coefficients were obtained. The **reprojection errors were within a pixel level accuracy** which is fairly acceptable.

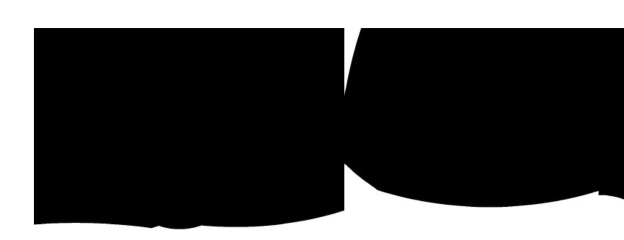




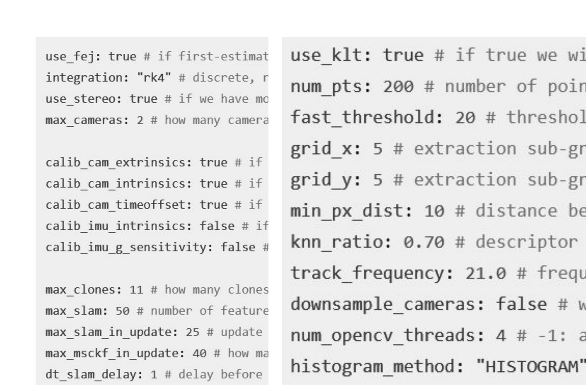
Rovioli and Openvins yielded the most accurate trajectory estimations, capturing the general shape of the initial trajectory effectively. In contrast, **Xivo's estimates were highly inaccurate**, exhibiting erratic behavior. **Orbslam3 failed to provide any estimates** due to its inability to find sufficient matching features across consecutive frames.

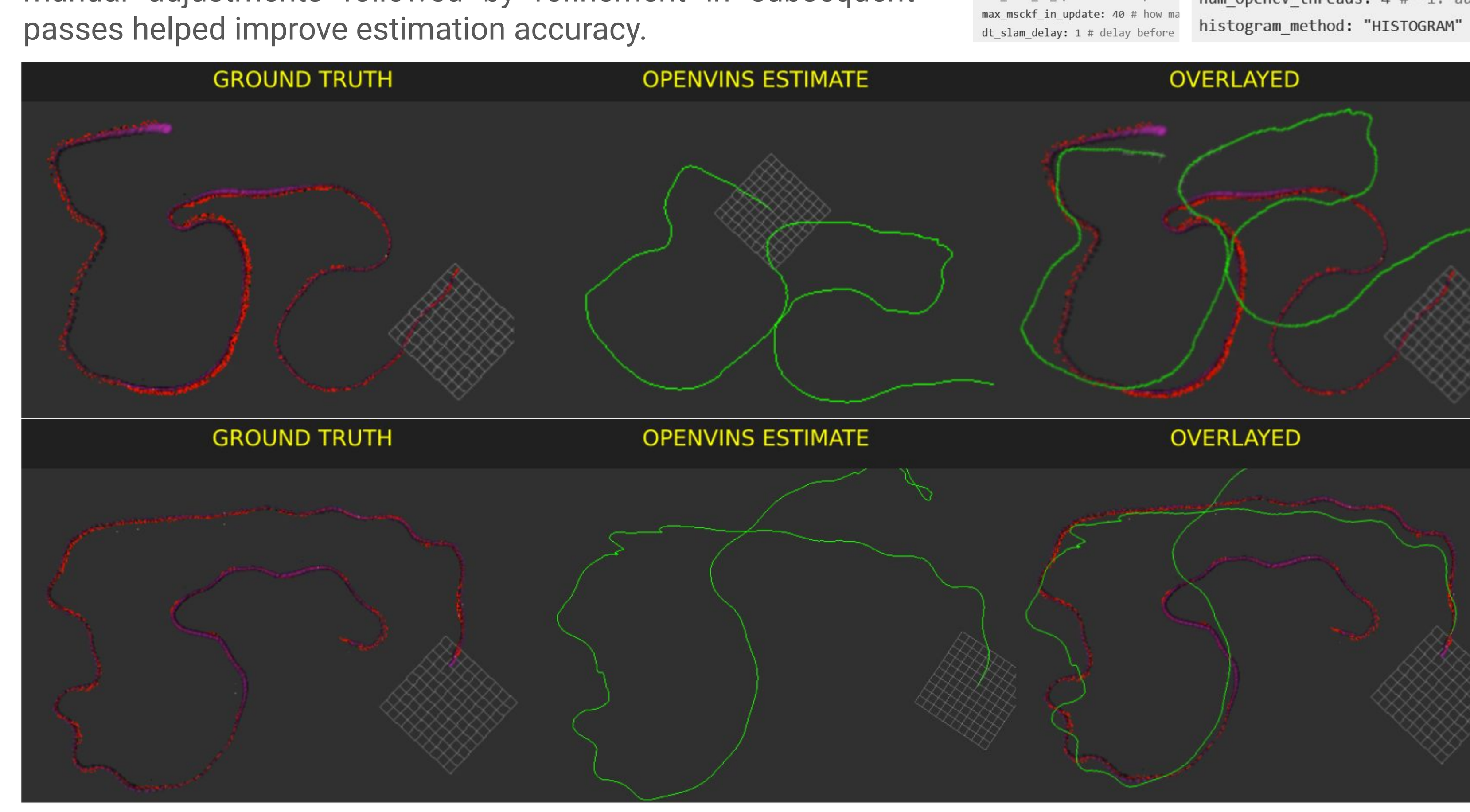
Several factors likely contributed to the **poor performance observed**, including **sensitivity to lighting** variations, **noisy compressed data** resulting from **erratic vehicle movement**.

Next step was enhancing estimation quality, starting with the crucial step of **validating the calibration**. Indoor dataset validation confirmed the correctness of intrinsic camera parameters and transformation matrices. Subsequent efforts involved **experimenting with image exposure, contrast**, and other enhancements, with **histogram equalization** showing promising results. However, attempts to improve tracking through **masking resulted in worsened outcomes**.



After this the algorithm parameters were tuned - adjusting parameters like the **number of tracked features** and **outlier rejection**. While testing all combinations was impractical, manual adjustments followed by refinement in subsequent passes helped improve estimation accuracy.

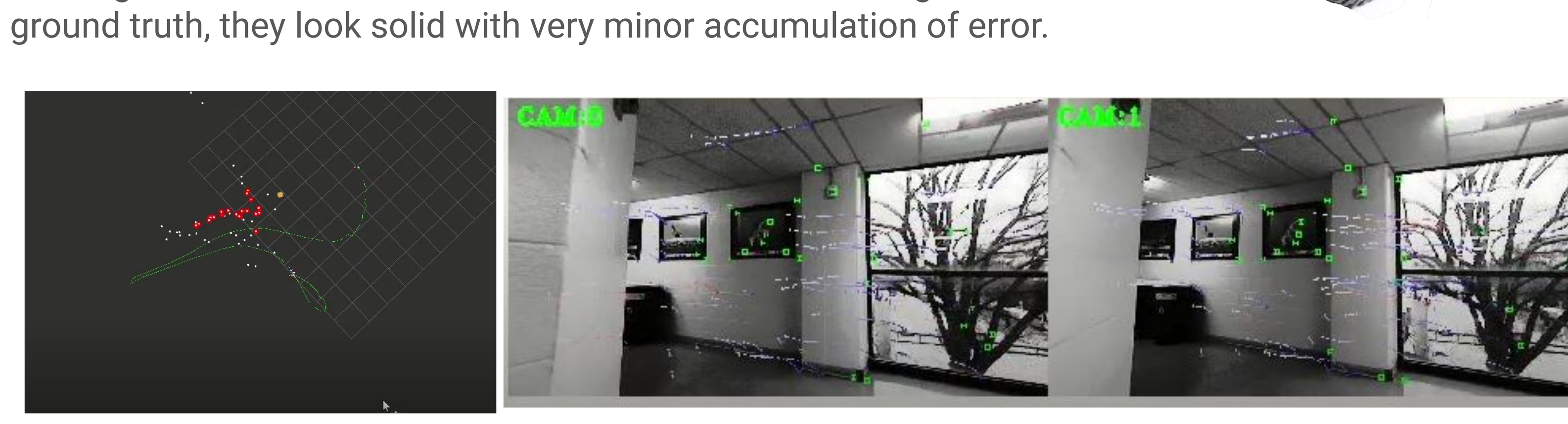




FUTURE WORK

New standard **MIT Racecar platform** will be used in the lab running on ROS2 Eloquent. Recalibration of hardware would be needed along with using new compatible libraries.

Work has already started and **intermediary results look promising**. Although, the results can't be verified due to not using GPS for ground truth, they look solid with very minor accumulation of error.



The final goal would be to **ensure loop closures** work properly ensuring that the accumulated drift error isn't too large when the car has to travel for a long time. This can result in navigation on shorter tracks in loops with a much higher confidence.

References:
1. [AutoRally] An Open Platform for Aggressive Autonomous Driving. Brian Goldfain, Paul Drews, Changxi You, Matthew Barulic, Orlin Velev, Panagiotis Tsiotras, James M. Rehg. Control Systems Magazine (CSM), 2019
2. [VIO] Y. Alkendi, L. Seneviratne and Y. Zweiri, "State of the Art in Vision-Based Localization Techniques for Autonomous Navigation Systems," in IEEE Access, vol. 9, pp. 76847-76874, 2021, doi: 10.1109/ACCESS. 2021.3082778.
3. [ORB-SLAM3] Carlos Campos, Richard Elvira, Juan J. G´omez, Rodr´ıguez, Jos´e M. M. Montiel and Juan D. Tard´os, ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM, IEEE Transactions on Robotics 37(6):1874-1890, Dec. 2021.
4. [OpenVins] Patrick Geneva, Kevin Eckenhoff, Woosik Lee, Yulin Yang, and Guoquan Huang, "OpenVINS: A Research Platform for Visual-Inertial Estimation," in Proc. of the IEEE International Conference on Robotics and Automation, Paris, France, 2020.
5. [Rovioli - Maplab] Michael Bloesch, Sammy Omari, Marco Hutter, and Roland Siegwart, "Robust Visual Inertial Odometry Using a Direct EKF-Based Approach," Conference Paper, 2015.
6. [XIVO] Xiaohan Fei and Stefano Soatto, "XIVO: An Open-Source Software for Visual-Inertial Odometry," 2019.
7. [Kalibr] Paul Furgale, T D Barfoot, G Sibley (2012). Continuous-Time Batch Estimation Using Temporal Basis Functions. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp. 2088–2095, St. Paul, MN.
8. [Allan Variance Method] https://upcommons.upc.edu/bitstream/handle/2117/103849/MScLeslieB.pdf?sequence=1&isAllowed=y