Study on the influence of the field-of-view on the precision of Monocular Visual SLAM

Aaryan Bhardwaj[1]
RAS (Systems Engineering)
Arizona State University
abhard27@asu.edu

Shivam Sharma[2]
RAS (Artificial Intelligence)
Arizona State University
sshar244@asu.edu

Walstan Baptista[3]
RAS (Artificial Intelligence)
Arizona State University
wbaptist@asu.edu

Team No.: 27

1 Introduction

Simultaneous Localization and Mapping (SLAM) is an essential technique for autonomous navigation, robotics, and augmented reality applications. Monocular Visual SLAM has become increasingly popular due to its low cost and flexibility. Field-of-View (FoV) is an important factor that influences the performance of Monocular Visual SLAM. A larger FoV provides more visual information, which can improve the precision of the estimated trajectory and the reconstructed environment.

In this study, we investigate the influence of FoV on the precision of Monocular Visual SLAM using ORBSLAM3 [1] with a perspective camera. Specifically, we aim to answer the research question: How does the FoV influence the precision of Monocular Visual SLAM using ORBSLAM3 with a perspective camera? Our findings can guide practitioners in selecting the appropriate FoV for their specific application requirements. Previous studies have investigated the influence of FoV on the precision of Visual-Odometry (VO) technology [2], which uses cameras to estimate the motion of a vehicle, but the influence of FoV on the precision of Monocular Visual SLAM using ORBSLAM3 with a perspective camera has not been fully explored.

Our study contributes to the existing literature on the influence of FoV on the precision of Monocular Visual SLAM by focusing on ORBSLAM3, which is a popular and widely used SLAM system. We used a Blender to render the camera's perspective as it traversed through an urban environment on a predetermined path. We then use ORBSLAM3 to get trajectory information from these images and compare that to our ground truth. Our results show that FoV has a significant influence on the precision of MonoSLAM using ORBSLAM3. A larger FoV improves the precision up to a certain point, beyond which the precision starts to degrade. Possible applications of this research include autonomous navigation in robotics, self-driving vehicles, augmented reality applications, and indoor mapping.

2 Related Work

Previous research has explored the influence of various camera parameters on the accuracy of visual odometry(VO) systems. However, there has been relatively little research on the impact of the FoV on the accuracy of visual SLAM, particularly in the context of large FoV cameras. The work presented in [2] is one of the few studies to explore this area. This paper evaluates the impact of camera FoV and optics on the quality of motion estimation in VO for urban canyon scenarios. While the choice of camera has a significant impact on the robustness and accuracy of VO algorithms, the best choice of FoV and optics is not straightforward, as increasing FoV while fixing the resolution decreases the angular resolution and lowers the measurement accuracy of a single camera pixel. The paper shows that smaller FoV cameras perform better in urban canyon scenarios. The experiments, which were conducted using both synthetic and real-world datasets, confirm the theoretical observations and provide an in-depth analysis of the challenges of adapting VO algorithms for large FoV cameras. The

36th Conference on Neural Information Processing Systems (NeurIPS 2022).

paper contributes to the existing body of work by providing insights into the impact of camera FoV and optics on VO performance in urban canyon environments.

ORB-SLAM3 [1] is a significant advancement in the field of simultaneous localization and mapping (SLAM), as it overcomes the limitations of previous systems by integrating visual SLAM with maximum a posteriori estimation, improving its robustness and accuracy in various environments. Its multiple map system allows it to survive periods of poor visual information and seamlessly merge previous maps when revisiting them. Additionally, the library's ability to reuse information from high parallax co-visible keyframes further boosts accuracy. Overall, ORB-SLAM3 is a powerful tool for visual-inertial and multisession SLAM, with impressive features that make it a leading choice in the field.

3 Baseline

To establish a baseline for our study, we conducted initial experiments using two different FoVs: 90 and 40 degrees. We used ORBSLAM3 with a perspective camera to estimate the camera trajectory and reconstructed environment for both FoVs. We measured the performance using Root Mean Square (RMS) error in translation, which is a commonly used metric for evaluating the precision of SLAM systems. Our baseline results show a significant impact of FoV on the precision of Monocular Visual SLAM. The RMS translation error was 4.5 meters for the FoV of 90 degrees, while it was 10 meters for the FoV of 40 degrees.

These results suggest that a larger FoV can improve the precision of Monocular Visual SLAM. The 90-degree FoV provided more visual information, leading to a more accurate estimation of the camera trajectory and reconstructed environment. In contrast, the 40-degree FoV provided less visual information, resulting in a higher RMS translation error. These findings are consistent with previous studies that have shown a positive correlation between FoV and the precision of Monocular Visual SLAM.

Our baseline results also highlight the importance of selecting an appropriate FoV for a given application. A smaller FoV may be sufficient for some applications that do not require high precision, but a larger FoV may be necessary for applications that demand more accuracy. Our study aims to further investigate this trade-off between FoV and precision using ORBSLAM3 with a perspective camera, which is widely used in robotics and augmented reality applications.

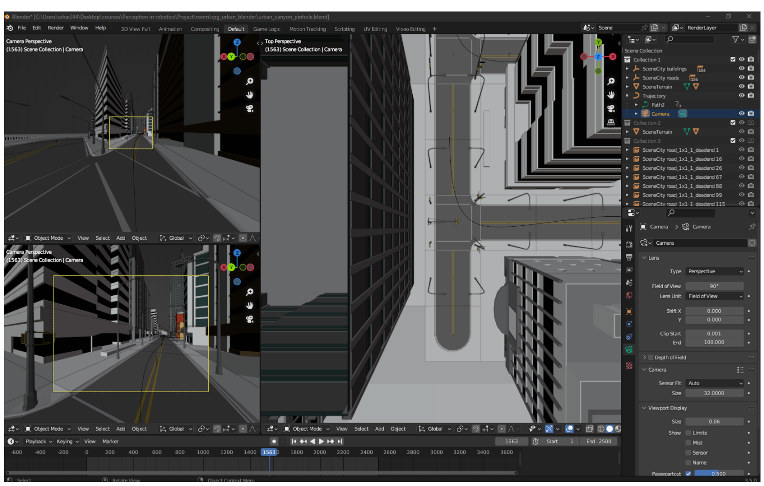


Figure 1: Blender Urban Canyon Environment

4 Approach

4.1 Dataset & Experimental Setup

We generated 2500 synthetic frames with various FoV ranging from 50 to 150 degrees using the pinhole cameras lens of fixed 480p resolution in the Urban Canyon Scene, in Blender given in [2] to simulate a forward-looking camera mounted on a car driving in a city environment.

In this study, we successfully set up ORB-SLAM3[5] on Docker using the method described in [3]. Subsequently, we installed the NVIDIA CUDA-based version after setting up the NVIDIA container toolkit. After obtaining the necessary user interface permissions, we initialized ORB-SLAM3 and tested it using the Monocular camera image feed from the EuRoC dataset. To generate the desired keyframe trajectories, we wrote camera calibration files for each test FoV. Calculating the focal length in pixel units for each FoV was a straightforward process; we used the focal length information from Blender and multiplied it by a factor of 20. Furthermore, we set up [4] in the container, to publish images from the generated frames directory to a channel that ORB-SLAM3 subscribed to. With this setup, we were able to generate keyframe trajectories and a point cloud map.

4.2 Experiments

The accuracy of 3D landmark measurements, which are crucial for geometric vision problems, depends heavily on feature correspondence in images. Therefore, we conducted an experiment to evaluate the accuracy of feature correspondence for perspective cameras with a constant image resolution. To investigate the impact of FoV on MonoSLAM using ORBSLAM3, we conducted experiments using different FoVs, with an interval of 10 degrees. We utilized frames of a robot moving in a blender environment 1 and input them to ORBSLAM3 2 to estimate the camera pose and the 3D map of the environment. We ensured consistency by utilizing the same environment and robot trajectory for all FoVs. We evaluated the accuracy of ORBSLAM3 by comparing the output trajectories with a ground truth trajectory.

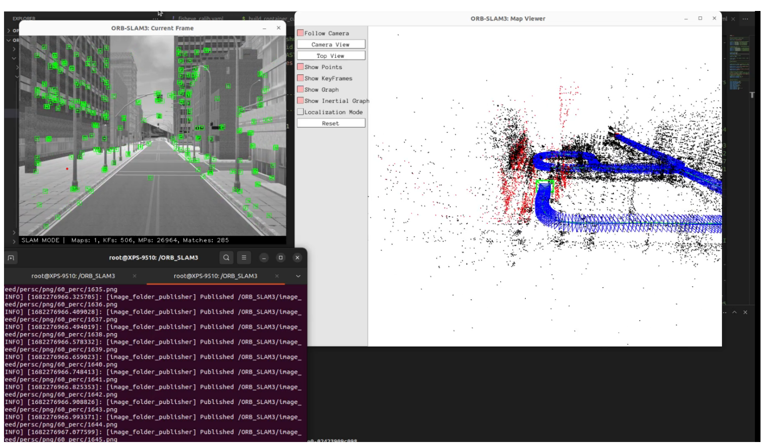
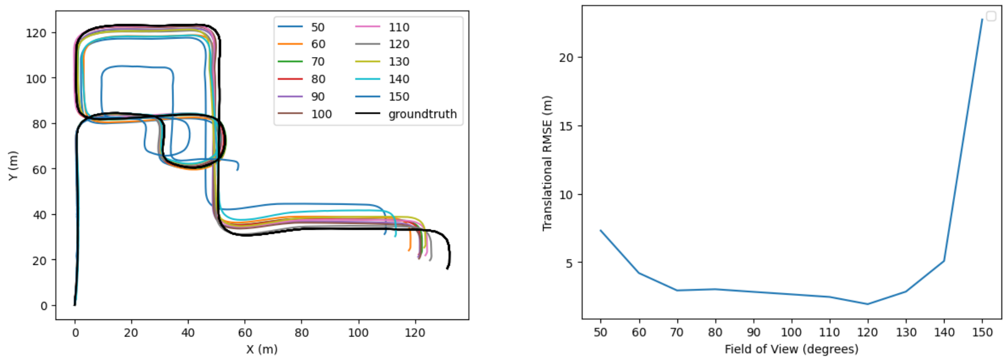


Figure 2: Visualization of ORBSLAM3 mapping output on a rendered environment

We calculated the RMSE between the estimated trajectory and the ground truth trajectory for multiple runs for each FoV. Our results indicate that the precision of MonoSLAM using ORBSLAM3 improves as the FoV decreases from 50 to 120 degrees, with the lowest RMSE attained at a FoV of 120 degrees. However, as the FoV increases beyond 120 degrees, the precision of MonoSLAM using ORBSLAM3 begins to degrade, with the lowest RMSE attained at a FoV of 120 degrees. In conclusion, our experiments demonstrate that FoV significantly affects the precision of MonoSLAM using ORBSLAM3, and a larger FoV enhances precision up to a certain point, beyond which the precision starts to degrade.

4.3 Results

Our experiments aimed to investigate the influence of FoV on the precision of Monocular Visual SLAM using ORBSLAM3 with a perspective camera. We rendered frames of a robot moving in a Blender environment at different FoVs and fed these to ORBSLAM3 and extracted estimated trajectories 3a.



(a) Trajectories vs Ground Truth for different FoV's

(b) FoV vs Translational RMSE

We evaluated the precision of the system using RMS error in translation for each FoV. Our results show a clear trend of the impact of FoV on the precision of Monocular Visual SLAM using ORBSLAM3 and a perspective camera 3b. We observe that increasing the FoV up to 120 degrees leads to a decrease in RMS error in translation. However, beyond this point, the precision starts to degrade with a FoV of 150 degrees resulting in a significantly higher RMS error in translation compared to the other FoVs. The FoVs between 70 to 130 degrees result in the best precision in terms of RMS error in translation, with 120 degrees having the lowest error.

Our findings are consistent with previous studies that have shown the influence of FoV on the precision of Monocular Visual SLAM. Our study expands upon these previous findings by using ORBSLAM3, a state-of-the-art SLAM system, and a perspective camera, which is commonly used in robotics and augmented reality applications. Overall, our experiments demonstrate the importance of selecting an appropriate FoV for achieving high precision in Monocular Visual SLAM applications using ORBSLAM3 and a perspective camera.

5 Conclusion & Future Work

In this study, we investigated the influence of FoV on the precision of Monocular Visual SLAM using ORBSLAM3 with a perspective camera. The results showed that the precision of the system was significantly impacted by the FoV, with the FoVs between 70 to 110 degrees resulting in the best precision in terms of RMS error in translation. Our findings are consistent with previous studies that have shown the importance of selecting an appropriate FoV for achieving high precision in Monocular Visual SLAM applications. However, our study expands upon these findings by using ORBSLAM3, a state-of-the-art SLAM system, and a perspective camera, which is commonly used in robotics and augmented reality applications. Our results demonstrate the significant impact of FoV on the precision of Monocular Visual SLAM using ORBSLAM3 and a perspective camera, highlighting the need for careful selection of FoV in such applications.

Possible applications of our research include robot navigation and mapping, autonomous driving, and augmented reality. By understanding the influence of FoV on the precision of Monocular Visual SLAM using ORBSLAM3 and a perspective camera, practitioners in these fields can make more informed decisions regarding the selection of FoV, resulting in higher precision and accuracy in their systems. Additionally, our study highlights the importance of continued development and refinement of SLAM systems and techniques to improve their performance in real-world applications.

References

[1] Campos, C., Elvira, R., Rodríguez, J. J. G., Montiel, J. M., & Tardós, J. D. (2021). Orb-slam3: An accurate open-source library for visual, visual–inertial, and multimap slam. IEEE Transactions on Robotics, 37(6), 1874-1890.

[2] Zhang, Z., Rebecq, H., Forster, C., & Scaramuzza, D. (2016, May). Benefit of large field-of-view cameras for visual odometry. In 2016 IEEE International Conference on Robotics and Automation (ICRA) (pp. 801-808). IEEE.

[3] https://github.com/jahaniam/orbslam3_docker

[4] https://github.com/amc-nu/RosImageFolderPublisher

[5] https://github.com/UZ-SLAMLab/ORB_SLAM3