# Performance Evaluation of Monocular Visual Odometry in Low-light Conditions and Narrow Field of View

Youn H. and Kim T. *
Dept. of Geoinformatic Engineering, Inha University, 100 Inha-ro, Michuhol-gu, Incheon, Republic of Korea
tezid@inha.ac.kr

## 1. Introduction

Monocular Visual Odometry (MVO) is a key component of autonomous driving technology, accurately determining a vehicle's relative position and orientation using a single camera. Recent research in the field of SLAM (Simultaneous Localization and Mapping) has been focusing on improving model robustness in challenging environments (Cadena et al.,2016). Particularly, efforts are being made to enhance adaptability to changing lighting conditions and to achieve consistent results across cameras with varying performance levels (Zhang et al, 2022). This study evaluates the performance of MVO in challenging environments, under low-light conditions and narrow FOV scenarios. In low-light environments, the quality of images captured by the camera deteriorates, making feature extraction and matching processes challenging. Additionally, as the camera's FOV narrows, the system's robustness to rotational movements tends to decrease. Considering these challenging conditions, this study evaluate the robustness of a monocular visual odometry system utilizing relative orientation techniques.

## 2. MATERIALS AND METHODS

### 2.1 Material

This study used publicly available data for Odometry/SLAM provided by the KITTI community. KITTI data is a representative benchmark for autonomous driving systems and computer vision research, collected in various road environments including urban, suburban, and highway settings. It allows for performance evaluation of autonomous driving systems under various traffic conditions.

### 2.2 Monocular Visual Odometry using relative orientation

In this study, we design a monocular VO based on relative orientation among adjacent video frames. Relative orientation in photogrammetry is a method of estimating the geometric relationship between two images. It calculates relative displacement and pose while fixing one of the three coordinate axes. The displacement represents movement in relative space, and from the geometric information obtained through this relative orientation, the movement vector and rotation estimate can be used in VO (Jung, 2018).

### 2.3 Monocular Visual Odometry Design

This section explains the monocular VO structure and processing sequence. The Shi-Tomasi corner feature algorithm is applied for real-time feature point extraction, and KLT (Kanade-Lucas-Tomasi) for feature point tracking. After feature point extraction and tracking, the relative orientation observation equation is established in images moving along the optical axis. We then calculate the geometric relationship between consecutive frames. When estimating image geometry using a single camera, it is impossible to estimate the actual scale. Therefore, additional information such as the camera height from the ground is essential to estimate the absolute scale. In this study, the absolute scale is estimated assuming that the height of the vehicle carrying the camera is known. At this time, as height differences occur in proportion to the actual moving speed, the RANSAC (Random Sample Consensus) algorithm is used to remove outliers, and finally, accurate position and pose are estimated.

### 2.4 Low-light Image Generate with CycleGAN

In this study, the KITTI dataset was transformed using the CycleGAN model to simulate low-light environments. By leveraging this model, realistic low-light images were generated through unsupervised learning. CycleGAN (Zhu et al., 2017) is a model capable of learning image-to-image translation between different domains without the need for direct pairing. It allows domain conversion without the need for precisely paired datasets required in supervised learning. Moreover, the core concept of CycleGAN, the cycle consistency loss, ensures the retention of important visual information by learning the process of converting an image to another domain and then back to its original domain. This mechanism allows for the simulation of lighting and contrast changes that may occur in real low-light environments while minimizing image distortion or information loss. Therefore, the KITTI dataset, a widely recognized dataset, was transformed to experiment in low-light conditions.

Figure 2. Original Image (Left) & Low-light Image (Right)

### 2.5 Evaluation in narrow FOV

In this study, the size of the KITTI data was modified to limit the FOV of the input images to evaluate MVO performance in narrow FOV environments. As the camera's FOV decreases, the instability of feature tracking increases, which negatively affects the overall accuracy of the system's position estimation. By cropping the horizontal and vertical FOV of the original KITTI data to a certain ratio, datasets with various levels of FOV restrictions were created. These datasets simulate environments like those with cameras using small sensors or lenses.

## 3. Results and Discussion

The performance evaluation of MVO in low-light environments and narrow FOV conditions was conducted using the proposed method. The results analyze the impact of these challenging conditions on the accuracy of MVO and evaluate the robustness of the system using relative orientation

### 3.1 Evaluation in Low-light condition

Figure 3 shows the comparison of the true trajectory and the estimated trajectory from the proposed method in original images(Left) and with low-light condition(Right). By comparing the trajectories in the figure below, it is evident that the localization performance of MVO in low-light environments has deteriorated. Despite conducting experiments in the same area, the transformed KITTI data using the CycleGAN model exhibited trajectory errors due to low-light conditions. The estimated trajectory(Right image) in low-light conditions (red line) showed greater positional errors compared to the normal lighting conditions and the ground truth trajectory (blue line). Nevertheless, in simple, straight paths, the difference between the two environments was minimal, and the errors were primarily observed in sections with rotations or sudden directional changes. This indicates that while feature extraction and matching may become challenging in low-light conditions, the system was still able to demonstrate localization performance in less complex paths.

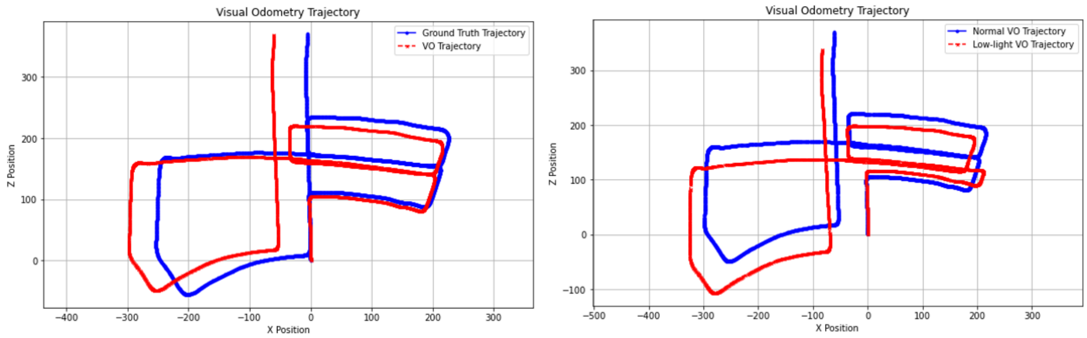


Figure 3. Comparison with Truth Trajectory (Left) & Original images Trajectory (Right)

### 3.2 Evaluation in narrow FOV condition

The results of the MVO system performance experiments in narrow FOV environments showed that as the FOV becomes more restricted, the overall accuracy of the system's position estimation tended to decrease. As the FOV decreased, the instability of feature tracking increased, leading to a decline in the overall position estimation accuracy of the system. Specifically, in situations involving rotational motion, the absolute shortage of feature point matches led to increased tracking errors. Table 1 presents a comparison of MVO system performance under various FOV conditions. The widest FOV of 1242x375 recorded the lowest RMSE (0.055444), indicating the best system performance. Conversely, as the FOV narrowed, there was a tendency for the RMSE values to increase. In particular, the FOV of 1000x250 showed the highest RMSE (0.632797), which highlighted the performance degradation when the FOV is reduced horizontally. Rotational errors increased nonlinearly as FOV narrowed, with higher RMSE observed at smaller horizontal FOVs. However, the relationship between errors was not always linear with respect to horizontal and vertical dimensions, and the extent of errors varied depending on the specific conditions. This variability may have been influenced by the shooting environment.

Table 1: Narrow FOV Experiments Results

| Size | RMSE | Rotation Error |
|---|---|---|
| 930x375 | 0.337576 | 0.00910 |
| 1000x375 | 0.254277 | 0.00973 |
| 1000x250 | 0.632797 | 0.00919 |
| 1050x280 | 0.429668 | 0.00980 |
| 1242x280 | 0.343162 | 0.01169 |
| 1242x300 | 0.172468 | 0.01178 |
| 1242x375 | 0.055444 | 0.01195 |

## 4. Conclusion and Recommendation

This study evaluated the performance of a relative orientation-based MVO system in low-light and narrow FOV environments. In low-light conditions, the degradation in the quality of images captured by the camera resulted in larger errors, particularly in sections involving rotations or sudden directional changes. In narrow FOV experiments, pixel-to-FOV conversion, a threshold of approximately 70 degrees (horizontal FOV) was identified .as the minimum requirement for maintaining accurate localization. When the horizontal FOV dropped below this threshold, system performance deteriorated sharply, with RMSE increasing significantly. Therefore, it is recommended that MVO systems be designed to ensure a minimum horizontal FOV of 70 degrees to maintain robust performance, especially in dynamic environments. The results of this study confirmed that the relative orientation-based MVO system could exhibit a certain level of robustness even in low-light and narrow FOV conditions. Future research should focus on improving performance in low-light environments by incorporating additional feature extraction and matching algorithms and addressing the inherent issue of rotational errors in MVO systems.

## References

Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., ... & 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.

Jung, J. (2018). Development of real-time driving recorder using affine and homography transformations from a single image (Master's thesis, Inha University). Incheon, South Korea.

**Reference to an article in online journals or online first [DOI]:**

Zhang, Y., Zhi, S., Lu, Z., & Lin, Z. (2022). Monocular vision SLAM research for parking environment with low light. *International Journal of Automotive Technology*, 23, 693-703. https://doi.org/10.1007/s12239-022-0063-5

Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. IEEE International Conference on Computer Vision (ICCV), 2242-2251. https://doi.org/10.1109/ICCV.2017.244

## Acknowledgments

This work is supported by the Korea Agency for Infrastructure TechnologyAdvancement(KAIA) grant funded by the Ministry of Land,Infrastructure andTransport(GrantRS-2022-00155763).