Robotics & Intelligent Systems Expo (RISEx 2025), Nov. 13-14, 2025, Edmonton, Canada

Egocentric geolocation estimation in the Farm context via Monocular Vision and Variable Baseline

Fardad Nasrollahzadehkhakiani, Mehran Mehrandezh, and Kambiz Ghaemi Osgouie
Faculty of Engineering and Applied Science, University of Regina, Regina, Canada.
Email: fnz657@uregina.ca

INTRODUCTION

Accurate geolocation is crucial for autonomous farming in agricultural settings. While stereo vision and LiDAR provide metric depth, they are costly and hardware-intensive [1, 2]. Monocular Depth Estimation (MDE) offers simplicity but faces issues with accuracy, scale ambiguity, and drift [3, 4]. This study proposes a novel MDE framework using a variable-baseline approach, where the first image serves as the reference frame, and subsequent images expand the triangulation baseline. Depth in the camera frame is combined with GPS data, which boasts a precision of 2 cm, through geometric transformation for absolute geolocation. Tested on a proprietary agricultural dataset, the method achieves a mean positioning error below 4%, surpassing traditional monocular odometry techniques. This scalable, cost-effective solution enables real-time adjustments for robust navigation in low-feature environments.

MATERIALS AND METHODS

The framework integrates four modules: (i) image distortion correction, (ii) feature extraction and pose estimation via ORB-SLAM2, (iii) triangulation of scene points, and (iv) global coordinate transformation.

A variable-baseline approach, depicted in Figure 1, anchors triangulation to the first image frame, expanding the baseline incrementally for enhanced depth accuracy. GPS data are fused with triangulated camera coordinates using an ellipsoidal Haversine formulation and East-North-Up (ENU) to Earth-Centred Earth-Fixed (ECEF) transformations to ensure precise mapping into latitude–longitude space. Experiments were performed on a proprietary agricultural dataset (1,857 images, 1280×720 pixels) under real farmland conditions.

RESULTS AND DISCUSSION

The variable baseline approach demonstrated superior performance compared to the fixed baseline method. As detailed in Table 1, the variable baseline attained a mean absolute error (MAE) of 0.3 % m, a root mean square error (RMSE) of 0.33% m, and a success rate of 94.7% (localization error < 0.5 m). In contrast, the fixed baseline triangulation method produced errors exceeding 4 %, with success rates falling below 20%. These results highlight the importance of reference-frame anchoring in reducing drift and scale inconsistencies in monocular systems. In comparison to conventional monocular SLAM methods, which typically report errors in the range of 1 to 2 m, the proposed approach achieves sub-meter accuracy without the need for additional sensors

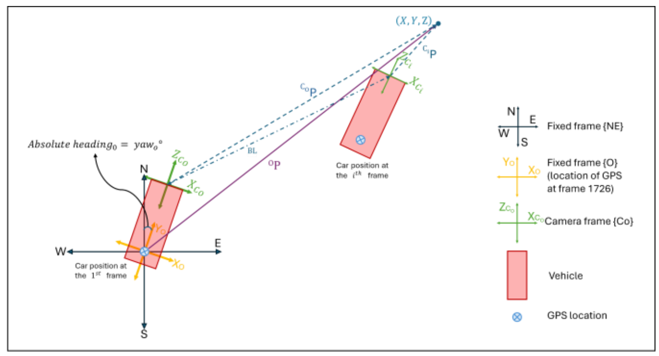

Fig 1. Variable-baseline approach: The first frame is fixed as the global reference origin, while subsequent frames I_i are incrementally incorporated for triangulation. This yields an expanding baseline that increases parallax over time, improving depth accuracy while preserving a consistent global origin.

CONCLUSIONS

This study demonstrates that monocular geolocation can achieve sub-meter accuracy when triangulation is consistently anchored to a fixed reference frame. The framework is cost-effective, scalable, and suitable for agricultural robotics where lightweight, single-camera solutions are preferred. Future work will incorporate transformer-based feature matching and sensor fusion to improve robustness under dynamic and low-texture conditions.

REFERENCES

[1] Kuo C-C, Chen L-K. *J Intell Technol Future Transp* **2378**: 50, 2025.
[2] Zhang J, Singh S. *J Robot Sci Syst* **2**: 1–9, 2014.
[3] Wang Z et al. *J Neural Process Lett* **55(7)**: 9743–9764, 2023.
[4] Forghani F et al. *J arXiv* **2503.15412**, 2025.

Table 1: Geolocation accuracy of fixed vs. variable baselines.

Method	Camera	MAE (%)	RMSE (%)	*Success Rate (%)*
Fixed baseline	Right	4.22	4.23	*17.9*
Variable baseline	Right	0.3	0.33	*94.7*