ISSN: 2832-790X
DOI: 10.33552/OJRAT.2026.04.000600
Online Journal of Robotics & Automation Technology
Iris Publishers

Mini Review

Copyright © All rights are reserved by John E Anderson

Cell Phone-Based Vision Terrain Reference Navigation

Jeffrey G Ruby[1], Jimmy R Carter[1], Melissa V Pham[1], William J, Shuart[2], Richard D Massaro[1], Robert L Fischer[1] and John E Anderson[2*]

[1]*US Army Geospatial Research Laboratory, Alexandria, USA*

[2]*US Army Geospatial Research Laboratory, Corbin Field Station, Woodford, USA*

***Corresponding author:** John E Anderson, USACE GRL Corbin, Woodford, Virginia USA

Received Date: November 06, 2026

Published Date: November 20, 2026

Abstract

Herein we review results and procedures of an effort to test a vision navigation workflow associated with the use of 3D terrain data and ground imagery as part of an application to estimate an accurate geographic position. Our approach demonstrates a passive technique hosted on a cell phone within an Android geospatial application as a plug-in we call P3R. P3R features a vision-based horizon-matching algorithm fused with adaptive processing to derive a position based upon both near- and far-field, fine and gross geographic terrain structures. Using a Samsung S23 cell phone's primary internal camera to acquire the necessary ground images and native compute power, our workflow featured both pre-processing and near-real-time processing of key elements for position estimations. Because the method does not rely on global positioning satellites, it is an alternative or companion technology implicated in autonomous or semi-autonomous systems that require precise, alternative navigation solution updates.

Keywords: Vision Navigation; Computer Vision; Horizon Matching; Edge detection

Introduction

The current state-of-the art in vision navigation is moving away from pre-mapped, rigid constructs (geo-centric) and into a world where systems can relate spatially between different data sets(feature-centric). Paw et al., (2026) [1] present new challenges in (algorithmic) computer vision where image feature selection is a key research area, particularly in switching between local and global data features to improve accuracy. Research trends are reported moving from data relational functions to fusion where perception is being introduced as part of image feature recognition [2].

As part of these trends, vision navigation is quickly maturing due to the accelerated development of algorithms for computer vision and electro-optic imaging coupled with improved processing hardware. These elements are driving a revival in algorithmic feature matching and passive and active visual odometry to achieve positioning. Many of these innovations have emerged from planetary exploration and autonomous navigation research and are incorporating new AI-based algorithms for positioning using feature-based strategies [3,4]. As vision navigation develops, it is also seen as a viable technology in the mitigation of the vulnerability (and reliability) of the global navigation satellite systems-GNSS [5]

Vision terrain reference navigation

Site Selection

Testing for our P3R application was performed at the National Oceanic and Atmospheric Administration's Corbin Field Station

This work is licensed under Creative Commons Attribution 4.0 License

in Woodford, Virginia, USA. The site features broad, open terrain vistas for acquiring ground imagery and reference control.

Hardware and Host Application

The testing scheme involved the use of the Samsung S23 cellular phone and main camera. Included on the S23: The Android Team Awareness Kit application (ATAK version 5.5), P3R (v.2.5.9-trx) and as pre-processed Numpy files housing 1-dimensional synthetic horizons (figure 1) generated at regular grid postings derived from 3D terrain data of the Corbin area (21 ha) and constructed using a ray tracing algorithm [6].

Image Acquisition

Sixty-six images collected at 22 ground control points (three images at different azimuths per point) were used to derive positions based on relating an extracted image horizon to the synthetic terrain horizons derived from the 3D data.



Figure 1: Grid post locations for synthetic horizon data base.



Figure 2: Phone image and 1D horizon extraction (in red) for matching.

Cell Phone Geo-positioning

Horizon extraction on the S23 utilizes a sky detection method incorporating a machine learning segmentation approach aided by an edge-detection algorithm. This workflow allowed for the horizon to be extracted in near-real time. To extract the horizon, the image is first segmented using the SegFormer transformer architecture [7], where the horizon boundary is defined by tracing the lowermost classified sky pixel in each vertical column. Leveraging the resulting semantic map, vertical occlusions such as light poles and flag poles are programmatically bypassed to prevent localized horizon distortion. Captured image horizons are then placed into the same cylindrical coordinate system as the stored synthetic horizons. With

azimuth angles stored as metadata for each synthetic horizon and the camera imagery providing roll, pitch, azimuth and pixel focal length (23.8 mm), a pixel horizon-to-elevation horizon transform can be generated. Finally, geolocation is determined for each post in the synthetic database using an iterative curve-matching process based on Mueen's Algorithm for Similarity Search (MASS) and Pearson's Correlation test for each image and database location [8]. For the sixty-six images collected, the average positional accuracy of 5.56 m was achieved (95% Conf.) with average RMSE of 3.15 m as compared to surveyed ground control based on real-time kinematic (RTK) position.

Conclusion

As a review of such a workflow housed in a limited compute environment, implications for autonomy and robotics are clear, particularly if attributes such as perception are introduced as part of the sensing strategy for location determination. While challenges remain, vision navigation technology presents a promising advantage for localization and navigation.

Acknowledgements

None.

Conflict of Interest

None.

References

1. Paw JKS, Tak YC, Kang LK, Benedict F (2026) Robotic vision system: Recent development trends and challenges. Terra Joule Journal 2: 158-163.
2. Khan J, Nayyer A, Ali Q, Mujahid M (2026) A comprehensive review of recent advancements in vision-and-language navigation. Discover Computing 29: 167.
3. Sarlin PE; DeTone D; Malisiewicz T, Rabinovich A (2020) SuperGlue: Learning Feature Matching with Graph Neural Networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, pp. 4938-4947.
4. Jiaqian H, Zhenhua J, Shuang L, Ming X (2024) Vision-aided inertial navigation for planetary landing without feature extraction and matching. Acta Astronaut 225: 316-327.
5. Zidan J, Adegoke E, Kampert E, Birrell S, Higgins M (2022) GNSS Vulnerabilities and Existing Solutions: A Review of the Literature. IEEE Access 9: 153960-153976.
6. Steger C, Steger B, Schar C (2022) HORAYZON v1.2: An efficient and flexible ray-tracing algorithm to compute horizon and sky view factor. Geosci. Model Dev 15: 6817-6840.
7. Xie E, Wang W, Yu Z, Anandkumar A, Alvarez J, Luo P (2021) SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers. Advances in Neural Information Processing Systems 34: 12077-12090.
8. Zhong S, Abdullah M (2024) MASS: Distance Profile of a Query Over a Time Series. Data Min Knowl Discov 38: 1466-1492.