

Research Article

Improving Multisensor Positioning of Land Vehicles with Integrated Visual Odometry for Next-Generation Self-Driving Cars

Muhammed Tahsin Rahman ¹, Tashfeen Karamat,²
Sidney Givigi,³ and Aboelmagd Noureldin ^{2,3}

¹Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada

²Department of Electrical and Computer Engineering, Queen's University, Kingston, ON, Canada

³Department of Electrical and Computer Engineering, RMCC, Kingston, ON, Canada

Correspondence should be addressed to Muhammed Tahsin Rahman; mt.rahman@mail.utoronto.ca

Received 19 January 2018; Accepted 26 March 2018; Published 2 May 2018

Academic Editor: Jose E. Naranjo

Copyright © 2018 Muhammed Tahsin Rahman et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For their complete realization, autonomous vehicles (AVs) fundamentally rely on the Global Navigation Satellite System (GNSS) to provide positioning and navigation information. However, in area such as urban cores, parking lots, and under dense foliage, which are all commonly frequented by AVs, GNSS signals suffer from blockage, interference, and multipath. These effects cause high levels of errors and long durations of service discontinuity that mar the performance of current systems. The prevalence of vision and low-cost inertial sensors provides an attractive opportunity to further increase the positioning and navigation accuracy in such GNSS-challenged environments. This paper presents enhancements to existing multisensor integration systems utilizing the inertial navigation system (INS) to aid in Visual Odometry (VO) outlier feature rejection. A scheme called Aided Visual Odometry (AVO) is developed and integrated with a high performance mechanization architecture utilizing vehicle motion and orientation sensors. The resulting solution exhibits improved state covariance convergence and navigation accuracy, while reducing computational complexity. Experimental verification of the proposed solution is illustrated through three real road trajectories, over two different land vehicles, and using two low-cost inertial measurement units (IMUs).

1. Introduction

As part of Intelligent Transport Systems (ITS), autonomous vehicles (AVs) are expected to become the norm in the near future; examples include self-driving taxis (implemented by companies such as nuTonomy in Singapore [1] and Uber), autonomous public services, and AV ride sharing. Such AV applications are predicted to reduce traffic congestions, improve passenger safety, and cut down overall operating costs [2]. However, implementing such a technology that would cause major societal changes comes with significant challenges, a major one being accurate vehicle positioning and localization.

Current systems rely heavily on the Global Navigation Satellite System (GNSS), such as the Global Positioning

System (GPS) and the Russian Globalnaya Navigazionnaya Sputnikovaya Sistema (GLONASS). These systems provide adequate accuracies when line of sight to four or more satellites is established. However, in situations where such line of sight is compromised, or when adverse multipath effects contaminate the received signals [3], or when the signals are spoofed or jammed [4], these systems are no longer sufficient as standalone solutions. Examples of such outage situations include urban centers, tunnels, parking lots, and dense foliage.

To bridge such outages, vehicular systems integrate GNSS with low-cost MEMS grade inertial measurement units (IMUs). Although being autonomous, a major drawback of systems employing these units is that they are unable to provide a stable navigation estimate during extended and

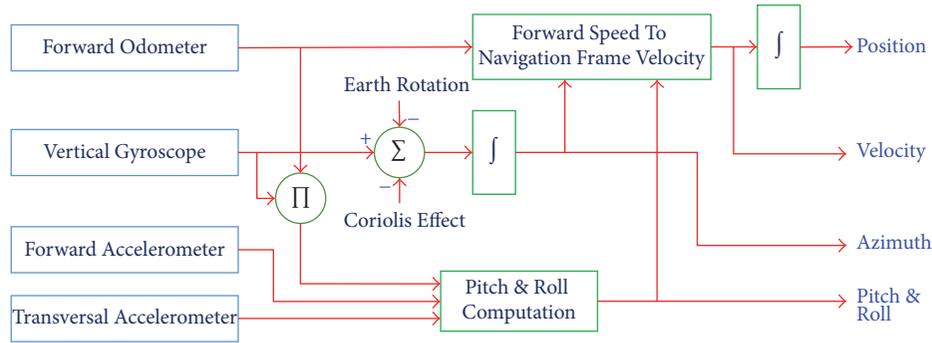


FIGURE 1: 3D-RISS block diagram. One integration is used for position and one for the azimuth angle.

frequent GNSS outages, situations common to AVs. Typically, these integration techniques are only able to tolerate GNSS outages that are less than two minutes in duration [5–7].

Vehicles are now coming equipped with cameras that are used in blind spot display, lane departure warning, and collision avoidance. The increasing availability of cameras provides an attractive opportunity to use vision based methods for vehicle localization, known in the literature as Visual Odometry (VO) [8]. However, VO based methods have several challenges. Unlike the inertial navigation system (INS) they are not autonomous, and hence nonstatic scenes in the camera field of view negatively affect the localization accuracy. Furthermore, they have high computational complexity, especially due to the inherent requirement of outlier rejection. The limitations of VO, although complimentary to the INS, cause the standalone solution to drift.

The diverse nature of each of these individual systems motivates the integrated INS/VO/GPS navigation solution. Essentially, this work is an extension of [9] and adds an enhanced VO solution into a multisensor framework for when GPS is unusable. This paper aims at the design and realization of an integrated multisensor positioning and navigation system capable of seamless and reliable positioning in all environments for land vehicles. The contributions of this paper are as follows:

- (1) Detailing a novel VO outlier rejection scheme, called R-AVO (Reduced-Aided Visual Odometry), to improve VO computational complexity and robustness to nonstatic scenes
- (2) Developing the INS/VO/GPS system architecture
- (3) Validating the proposed algorithms over several real road trajectories, utilizing different grade IMUs and different land vehicles.

In the following, Section 2 provides an overview of INS algorithms, VO advances, and recent work on their integration. Section 3 then details the methodology used in this work, including the integration scheme. Section 4 describes the experimental setup and the trajectories analyzed, which is followed by a discussion of using the proposed solution as compared to the traditional INS/GPS solution. This section also evaluates the algorithms' performances. Finally, the paper concludes with Section 5.

2. Related Work

2.1. Inertial Navigation System (INS). In a strap-down INS, 3-axis orthogonal gyroscopes are used to update the transformation between the vehicle body frame and the navigation frame. Then, accelerometer readings pass through this transformation and provide a navigation solution. As shown in [10], exploiting nonholonomic constraints of a land vehicle allows for the use of 3-axis orthogonal gyroscopes with a single forward accelerometer.

To further simplify the system, authors in [11] suggest using a single gyroscope, two accelerometers, and a vehicle odometer reading forward speed. The system block diagram is shown in Figure 1. Given that the vehicle plane remains mostly horizontal, this method of mechanization (called 3D Reduced Inertial Sensor System (3D-RISS)) has been shown to outperform the other methods [12].

This mechanization algorithm achieves improved accuracy for several reasons, which are fully discussed in [12, 13]. Most notably, using a single gyroscope in attitude estimation means that only one integration needs to be performed to determine the INS misalignment. As is well studied [7], gyroscopes exhibit stochastic bias, and integrating the sensor output introduces drift to the overall rotation matrix that maps the body frame to the navigation frame. Over time, three drifts from three gyroscopes very quickly lead to an unusable rotation matrix estimate. A single up-facing gyroscope means that this drift is limited to only the azimuth angle computation. Since pitch and roll are directly calculated using accelerometer and odometer readings they exhibit no drift.

2.2. Visual Odometry (VO). Visual Odometry, a term coined in [8], involves estimating a dead-reckoning navigation solution based on a sequence of images. A block diagram of the VO method is shown in Figure 2.

2.2.1. Feature Detection and Matching. Significant features in the images, also called keypoints, are detected in one frame and then matched in the next frame or frames. Authors in [14] evaluate detector-descriptor combinations that perform best for a moving camera. Most notably, Harris, Shi-Tomasi, and FAST detectors work best for shorter baseline distances without perspective distortion, which is characteristic of AVs, especially when the framerate is greater than about 5 Hz.

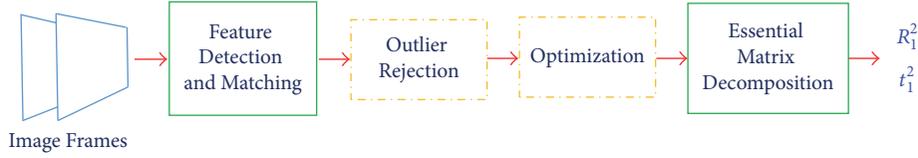


FIGURE 2: The generic VO system pipeline. Green boxes represent noniterative algorithms while yellow dash-dot boxes represent iterative algorithms.

Although the FAST detector is the most efficient, it suffers from lower repeatability and poorer keypoint registration accuracy, while Harris and Shi-Tomasi detectors have similar reported reprojection errors [14].

One of the main limitations of VO is the drop in feature detection and matching performance when operating in degraded visual environments (DVE). Previous work in this area include studies that analyze the system performance in low light conditions [15], algorithms that exploit color information [16], filters that aim to eliminate rain and snow from individual images [17], and localized image sequence matching in SLAM methods [18]. However, within the scope of this study, it is assumed that the effects of DVE are minimal.

2.2.2. Motion Estimation. A relation between matched key points in two frames is established through the epipolar constraint [19]: $\hat{p}^T E \tilde{p} = 0$, where \hat{p} represents the normalized key points in the second frame, \tilde{p} represents the normalized key points in the first frame, and E represents the *Essential Matrix*. The two most common algorithms used in solving for E are the 8-point algorithm [20] and the 5-point algorithm [21], so called because of the minimum number of key point matches needed to instantiate a 6-DoF solution.

More recently, nonholonomic constraints of planar vehicular motion have been exploited, giving rise to the 2-point algorithm [22] and the 1-point algorithm [23], both of which solve for a 3-DoF solution. Due to the constrained motion and the reduced minimum number of points needed, these algorithms are more robust to outliers and have lower computational complexity. However, the 2-point algorithm was tested only for a slow moving indoor robot and encountered completely erroneous solutions when the first randomly chosen set of points were incorrectly matched. As for the 1-point algorithm, it is based on the Ackermann steering principle and makes the assumption that the camera is placed on top of, or at least close to, the rear wheel axis. This assumption did not hold true for one of the test vehicles used in this project (shown later in Figure 5(a)). As such, and to keep the proposed system generic, this paper uses the 5-point algorithm.

2.2.3. Outlier Rejection. One of the major motivations for reducing the number of points needed to instantiate a solution for the *Essential Matrix* is the processing time needed for removing outliers. Random sample consensus (RANSAC) [24] is the most common method in this regard.

Another widely used method that lends itself to VO is the maximum likelihood estimation sample consensus (MLESAC) [25], which is an extension of m-estimator sample

consensus (MSAC) [26], which in turn is a conceptual generalization of RANSAC. MLESAC calculates the likelihood that a given hypothesis is true. It does this by representing the error distribution of the residuals as a *mixture model*, referring to the mixture of inlier and outlier error residual distribution probabilities, and then finds the hypothesis that maximizes this probability. The error distribution depends on the *mixing parameter*, which is the probability that a given point is an inlier. Authors in [27] use a priori information about the *mixing parameter* to improve the overall solution. Other schemes, such as [28, 29], suggest novel methods for building on the MLESAC scheme to improve accuracy and computational efficiency.

2.3. Sensor Integration Techniques. Much previous work on integrating VO in robotic and vehicular applications fuses a 6-DoF IMU with the processed information available at the end of the VO pipeline; see, for example, [30–33]. Authors in [34] present temporal aliasing constraints on the implementation of VO, showing how inertial sensors can improve the computational requirements of the overall system. A straightforward implementation of INS/VO/GPS integration is to use a switching architecture, where a sensor health monitoring system classifies available sensor readings and utilizes them appropriately. Such a system is necessarily a modular one, loosely integrating additional information based on the information quality [35, 36]. This is an attractive architecture for current multisensor systems due to the low computational overhead. Given the emerging trend of using LiDAR, radar, multiple IMUs, and multiple cameras in AVs, this paper limits its discussion to loosely coupled integration.

3. Materials and Methods

Several inherent advantages of 3D-RISS mechanization can be used to improve the VO navigation estimate before being fed into the filter. Authors in [37, 38] integrated information from other sensors to improve VO at the feature matching stage. While this would, conceptually, lead to better outlier rejection, and hence improved VO performance, a variant of it is already inherent in the KLT tracker algorithm. Furthermore, classic outlier rejection methods would still need to be applied to account for closely spaced outlier matches. A more natural implementation of multisensor integration before the final VO output stage is to directly aid outlier rejection schemes. Since the main computation bottleneck is outlier rejection, different sensors can be used to alleviate the iterative nature of this stage. Termed R-AVO (3D RISS-Aided Visual Odometry), this aid includes two major parts:

improving the outlier rejection step and resolving scale factor ambiguity.

3.1. Reduced-Aided Visual Odometry (R-AVO). In the following notations, the subscript (mech or cam) represents the system that generates the particular solution; the superscript (mech or cam) represents the coordinate frame of the solution. For example, $R_{1,\text{mech}}^{2,\text{cam}}$ is the rotation matrix from epoch 1 to 2, generated through the mechanization algorithm and expressed in the camera coordinate frame.

The R-AVO outlier rejection algorithm works as follows:

- (1) Compute the 3D-RISS mechanization solution.
- (2) Convert the INS solution to the camera coordinate frame and then to the *Essential Matrix* form $E_{\text{mech}}^{\text{cam}} = [t]_{1,\text{mech}}^{2,\text{cam}} R_{1,\text{mech}}^{2,\text{cam}}$.
- (3) Run VO feature detection and tracking.
- (4) Using the detected features and $E_{\text{mech}}^{\text{cam}}$ calculate the expected position of the matched features.
- (5) Find the geometric distance between the expected feature positions and matched feature positions.
- (6) Based on a threshold decide which features are inliers

Using 3D-RISS, the changes in the computed Euler angles, pitch, roll, and azimuth, from one epoch to the next are used to generate the rotation matrix $R_{1,\text{mech}}^{2,\text{mech}}$. The odometer speed multiplied by the time increment Δt is then projected through this transformation to get the incremental translation in the body frame.

$$t_{1,\text{mech}}^{2,\text{mech}} = \begin{bmatrix} t_{\text{right}} \\ t_{\text{forward}} \\ t_{\text{up}} \end{bmatrix} = v_{\text{od}} \Delta t \begin{bmatrix} \sin(\Delta A) \cos(\Delta\theta) \\ \cos(\Delta A) \cos(\Delta\theta) \\ \sin(\Delta\theta) \end{bmatrix}, \quad (1)$$

where

$t_{1,\text{mech}}^{2,\text{mech}}$ is incremental 3D-RISS based mechanization frame translation vector;

Δt is incremental time between epochs;

ΔA is incremental azimuth angle;

$\Delta\theta$ is incremental pitch angle.

Then, the incremental rotation and translation of the vehicle in the camera frame are given by

$$\begin{aligned} R_{1,\text{mech}}^{2,\text{cam}} &= \Omega_{\text{mech}}^{\text{cam}} R_{1,\text{mech}}^{2,\text{mech}}, \\ t_{1,\text{mech}}^{2,\text{cam}} &= \Omega_{\text{mech}}^{\text{cam}} t_{1,\text{mech}}^{2,\text{mech}}, \end{aligned} \quad (2)$$

where

$\Omega_{\text{mech}}^{\text{cam}}$ is coordinate transformation from mechanization frame to camera frame;

$R_{1,\text{mech}}^{2,\text{cam}}$ is incremental 3D-RISS based rotation matrix in camera frame;

$R_{1,\text{mech}}^{2,\text{mech}}$ is incremental 3D-RISS based rotation matrix in mechanization frame;

$t_{1,\text{mech}}^{2,\text{cam}}$ is incremental 3D-RISS based translation vector in camera frame;

$t_{1,\text{mech}}^{2,\text{mech}}$ is incremental 3D-RISS based translation vector in mechanization frame.

Sensor misalignment [39], appearing as errors in the coordinate transformation matrix $\Omega_{\text{mech}}^{\text{cam}}$, could lead to misclassification of inlier and outlier features. Once the above are determined, the translation vector is scaled to unit length, changed to its skew symmetric form, and used to determine the corresponding *Essential Matrix* $E_{\text{mech}}^{\text{cam}}$.

3.2. Geometric Distance. Well-established metrics for distance computation can be implemented to find the error between 3D-RISS based expected feature positions and the feature matching algorithm. Features within a threshold will then be taken as inliers and features outside the threshold will be outliers.

The Gold Standard algorithm [19] is the yardstick to which other algorithms are compared. In the context of R-AVO, this algorithm first triangulates the matched features from both frames into 3D feature coordinates using $E_{\text{mech}}^{\text{cam}}$. The 3D world points are then reprojected back onto the frames using the camera matrices, and the distances between the reprojected points and the original features are computed. Each feature, therefore, has a corresponding distance error associated with it, which is based on the features conformance to the 3D-RISS algorithm.

Computationally efficient approximations that do not involve triangulation, such as the Sampson error, could also be considered. The Sampson error is given by

$$d_s = \frac{(\tilde{p}^T F \tilde{p})^2}{(F \tilde{p})_1^2 + (F \tilde{p})_2^2 + (F \hat{p})_1^2 + (F \hat{p})_2^2}, \quad (3)$$

where

F is fundamental matrix;

\hat{p} is pixel coordinates in second frame;

\tilde{p} is pixel coordinates in first frame.

Although this method has led to the best approximation in the context of multiple view geometry, its performance significantly worsens when integrated with 3D-RISS. An explanation for this performance degradation is that geometric feature errors are assumed to remain small between different views, typically within 1 pixel, which does not hold when comparing expected and matched features from two completely different systems. A comparison of the inlier feature set for MLESAC, R-AVO-G (using the Gold Standard), and R-AVO-S (using the Sampson error) is shown in Figure 3.

An interesting point of note is that using R-AVO provides a certain degree of robustness to nonstatic elements in the scene. Since only the vehicle dynamics are used, features corresponding to scene motion rather than vehicle motion

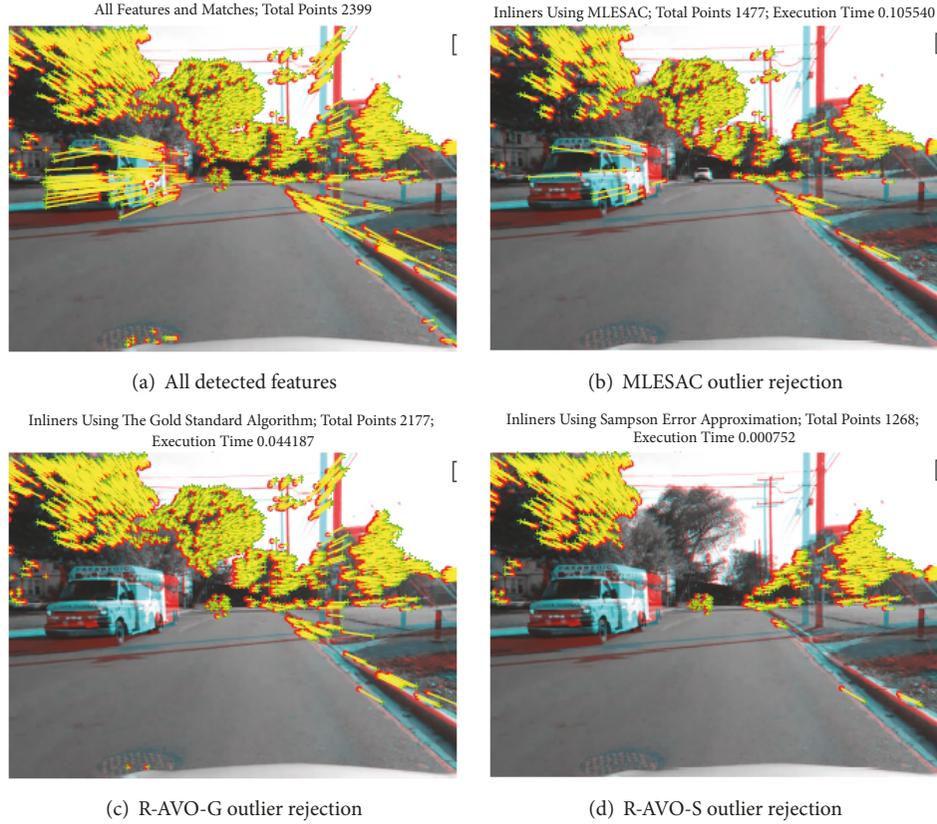


FIGURE 3: Comparison of feature inliers using different algorithms for outlier rejection. 2399 points were detected overall. MLESAC in (b) had 1477 inliers and took 0.106 s. R-AVO-G in (c) had 2177 inliers and took 0.044 s. R-AVO-S in (d) had 1268 inliers and took 0.0008 s.

are automatically rejected. This is apparent by the fact that features on the moving ambulance in Figure 3, which do not faithfully represent the camera motion, were rejected. The limitation of the Sampson error metric is also evident. Features on the tree in the background were determined as outliers, even though MLESAC and the Gold Standard algorithm considered them inliers.

3.3. Scale Factor Ambiguity Resolution. A modest, yet important, addition to the INS/VO solution is using the odometer to compensate for the scale factor ambiguity inherent in VO. Although the scaled translation vector $t_{1,mech}^{2,cam}$ is used to triangulate the 3D world points in R-AVO-G, the scale is lost at the final stage when computing motion from 2D to 2D correspondences. This scale must, therefore, be reintroduced at the end of the VO pipeline. The final translation vector is the unit translation vector from VO multiplied by $v_{od}\Delta t$.

As will be seen next, the scale factor of the odometer is considered in the integration filter, which directly relates to the scale factor of the VO solution. In effect, estimating errors in the odometer leads to compensating for a significant source of error in monocular VO.

3.4. Integration Filter Architecture. This paper integrates the 3D-RISS solution with 5-point VO using an indirect extended Kalman filter (EKF). The nonlinear motion model must be linearized to be used in the EKF, which is done using a

Taylor series expansion and ignoring the higher order terms. For brevity, these are not presented here; however, interested readers may refer to [40] for an in-depth derivation. The error state vector $\delta\mathbf{x}$ is given by

$$\delta\mathbf{x} = [\delta\varphi \ \delta\lambda \ \delta h \ \delta v_e \ \delta v_n \ \delta v_u \ \delta A \ \delta a_{od} \ \delta b_z]^T, \quad (4)$$

where

$\delta\varphi$ is error in latitude;

$\delta\lambda$ is error in longitude;

δh is error in height;

δv_e is error in east velocity;

δv_n is error in north velocity;

δv_u is error in up velocity;

δA is error in azimuth;

δa_{od} is stochastic odometer scale factor error;

δb_z is stochastic gyroscope drift.

Since the pitch and roll angles are computed *without integration* in 3D-RISS using the accelerometers and the up-facing gyroscope measurement, they do not experience drift. Furthermore, the pitch and roll angles from VO are prone to noise [35] and are inherently less accurate than the 3D-RISS solution, especially since the accelerometer data

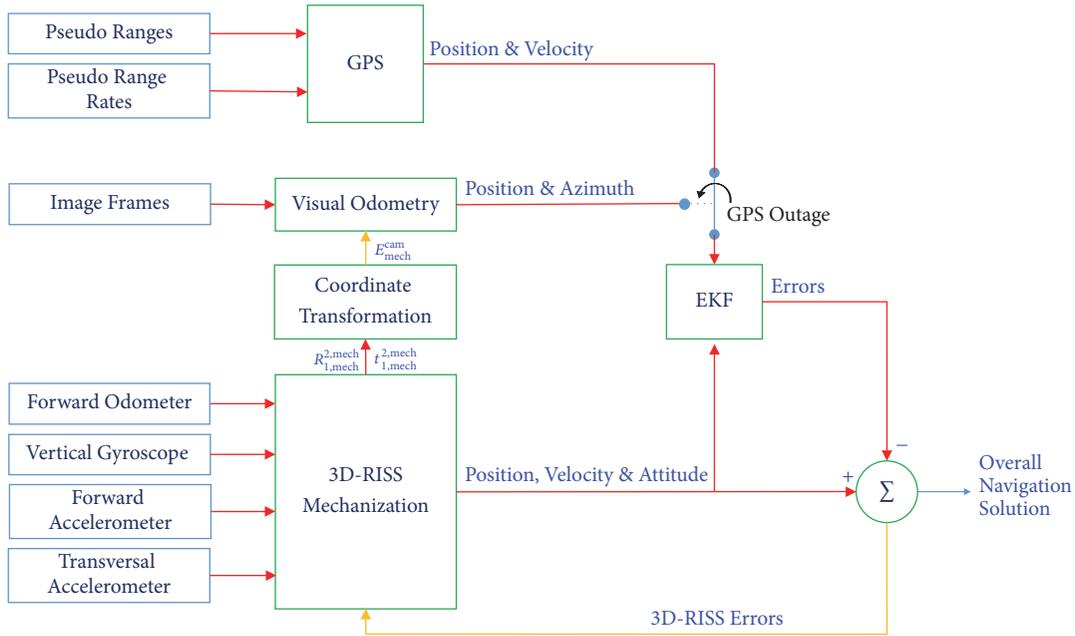


FIGURE 4: INS/VO/GPS integration using a switching architecture.

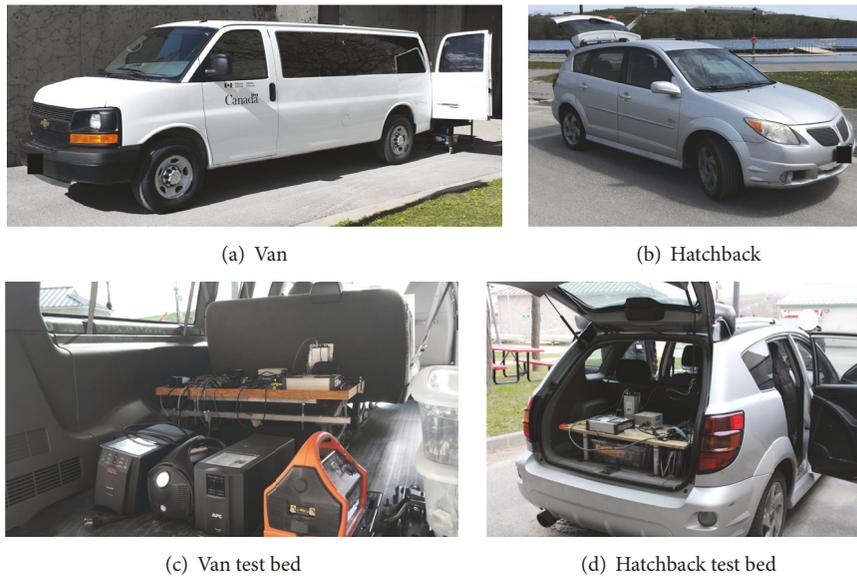


FIGURE 5: The vehicles used in the trajectories and the sensor test bed rigidly mounted.

is first denoised. The azimuth angle, however, requires an integration, and since land vehicle navigation is sensitive to azimuth angle errors [41], VO is used to suppress the resulting drift. Therefore, in the error state formulation, only the error in azimuth angle δA and the gyroscope bias δb_z are considered.

δa_{od} and δb_z are the parameters that model the odometer scale factor errors and the gyroscope drift based on a first-order Gauss-Markov process [7].

The system block diagram is shown in Figure 4. Different states are observed by the two measurement systems, so the

design matrix H will vary based on the active system. During GPS availability, the design matrix is

$$H_{GPS} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}. \quad (5)$$

TABLE 1: Comparison of different IMUs.

Specification	SPAN-CPT	Crossbow	VTI
Update rate	100 Hz	100 Hz	20 Hz
Gyroscopes			
Bias	20 deg/h	2 deg/s	1.5 deg/s
Scale factor	0.1%	1%	2%
Angle random walk	0.0667 deg/ $\sqrt{\text{hr}}$	2.22 deg/ $\sqrt{\text{hr}}$	0.86 deg/ $\sqrt{\text{hr}}$
Accelerometers			
Bias	50 mg	30 mg	70 mg
Scale factor	0.4%	1%	1%
Range	10 g	2 g	6 g

When there is a GPS outage, the design matrix is

$$H_{\text{VO}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \quad (6)$$

Similar to when using GPS as the measurement update, the VO measurement noise matrix $R_{k,\text{VO}}$ needs to be determined at each epoch. This parameter was set to a scaled version of the average geometric distance for the inliers; see Section 3.2.

4. Results and Discussion

Abbreviations for the different algorithms experimentally compared in this work are listed below. In all cases, the mechanization was done using 3D-RISS:

- (i) (INS)/(GPS): classic INS GPS integration
- (ii) (INS)/(MLESAC-VO)/(GPS): switching architecture with VO updates, using MLESAC based outlier rejection
- (iii) (INS)/(R-AVO-G)/(GPS): switching architecture with VO updates, using R-AVO based outlier rejection with the Gold Standard distance metric
- (iv) (INS)/(R-AVO-S)/(GPS): switching architecture with VO updates, using R-AVO based outlier rejection with the Sampson error distance metric.

4.1. Equipment. The overall setup consisted of an external test bed mounted onto a vehicle. Two vehicles were used: a GMC Savana van shown in Figure 5(a) and a Pontiac Vibe hatchback shown in Figure 5(b). For the van, the last row of seats were removed and the test bed was rigidly fixed in its place, attached to the vehicle floor. This is shown in Figure 5(c). With the hatchback, the cargo was modified with floor hooks to which the testbed was rigidly fixed, as shown in Figure 5(d). A MiVue 388 video recorder camera was attached to the outer windshield of each vehicle. Although the camera captured video at 30 fps, the VO algorithm was run at 5 Hz, owing to the increased computational demand of higher update rates; the VO localization estimates were then



FIGURE 6: The different sensors used in the experiment, shown with the respective model numbers.

concatenated and used to update the RISS estimate at 1 Hz, similar to GPS. This method allowed for consistency at the update stage of the filter. The vehicles' OBD-II port was used to record forward speed.

The ground-truth solution, also called the reference solution, consisted of a Novatel SPAN unit, which provides a tightly coupled INS/GPS integrated solution. This unit included an OEM4 GPS receiver integrated with a tactical-grade SPAN-CPT IMU (model OM-20000122). The other two low-cost IMU sensors that were used to validate the algorithms were the MEMS grade Crossbow IMU (model IMU300CC-100) and the MEMS grade VTI IMU (model SCC1300-D04). More details of the three IMUs are given in Table 1. Since the primary concern is the algorithm performance during outages, the GPS receiver used was the same for all algorithms. The test bed with these sensors marked on it is shown in Figure 6.

4.2. Performance Metric. The metric used to quantify the algorithm performance is the 2D root mean square (RMS) error. Since the context is vehicular motion, the metric is limited to 2D, even though all algorithms provide the full 3D solutions. The 2D RMS error is computed as follows:

$$2\text{DRMSError} = \sqrt{\frac{1}{K} \sum_k (2\text{Dpos}_{k,\text{ref}} - 2\text{Dpos}_{k,\text{ekf}})^2}. \quad (7)$$

To better visualize the results, trajectory portions are overlaid on maps. These were created using the freely available online tool at <http://www.gpsvisualizer.com/> [42]. In all cases, color codes are as follows:

- Novatel reference: red
- Outage portions: black

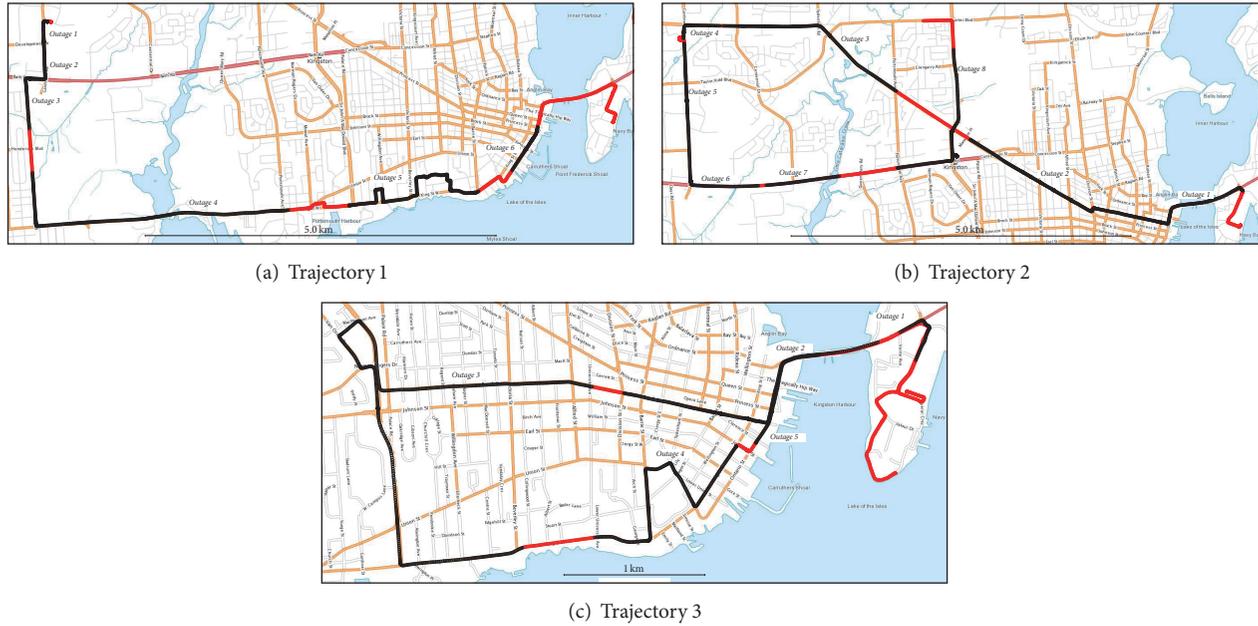


FIGURE 7: Trajectory and outage portions of the three experimental tests.

(INS)/(GPS): green

(INS)/(MLESAC-VO)/(GPS): magenta

(INS)/(R-AVO-G)/(GPS): cyan

(INS)/(R-AVO-S)/(GPS): amber.

Data was analyzed using Matlab 2017a, on an x64-based processor, Intel i7-4720HQ CPU @ 2.6 GHz. The PC RAM was 16 GB. Computation times for (INS)/(MLESAC-VO)/(GPS), (INS)/(R-AVO-S)/(GPS), and (INS)/(R-AVO-G)/(GPS) were determined for each solution epoch in which an outage occurred. Given that the time taken for the INS mechanization was negligible, (INS)/(GPS) times alone have not been recorded.

4.3. Trajectories. Three real road trajectories were conducted in Kingston, Ontario, Canada. During the trajectory, outages were simulated in postprocessing to reflect situations commonly encountered by AVs. The first two trajectories used the long GMC Savana van, shown in Figure 5(a). The first trajectory lasted 34 minutes, consisted of six simulated outages, and tested the VTI IMU. The second trajectory lasted 43 minutes, consisted of 8 simulated outages, and tested the Crossbow IMU. The third trajectory used the older Pontiac Vibe hatchback, lasted 33 minutes, consisted of five simulated outages, and tested the VTI and Crossbow IMUs together. The trajectories are shown in Figure 7.

4.4. Navigation Accuracy. For each of the three trajectories, the 2D RMS errors during the outages are shown in Figure 8, with some trajectories overlaid on maps in Figure 9. In most cases, the inclusion of VO, apart from just eliminating error spikes, causes significant accuracy improvement. The proposed R-AVO paradigm further improves the INS/VO/GPS

solution when the Gold Standard algorithm is used to determine outliers. As explained in Section 3.2, the Sampson error metric's approximation does not faithfully represent the underlying sensor dynamics, resulting in erroneous feature matches that cause errors in the VO solution.

There are several cases, however, where the accuracy of classic INS/GPS is similar to INS/VO/GPS. These are mainly for Outage 1 and sometimes Outage 2 across the trajectories. An explanation for this difference from the norm is that the filter had a lot of time to converge before the first outage, trusting its estimated bias for the INS. So when the outage does occur, the system continues to trust the INS more than the VO solution, leading to similar accuracies for the INS/GPS and the INS/VO/GPS cases.

When comparing the different types of VO algorithms used, a point of note is that in Trajectory 3 the accuracy of R-AVO-G was similar to or only slightly better than MLESAC based VO. This can be attributed to the fact that the vehicle used was different. Given that it was an older and smaller land vehicle, the resistance to road anomalies was less, and hence the vibrations experienced by the IMU were greater [43]. This can be seen by first comparing the trajectories' performance during GPS availability (and hence without VO), shown in Table 2. The filter implemented using the VTI IMU outperformed the Xbow IMU, and the GMC van outperformed the Pontiac hatchback. Given that the same GPS receiver was used, the degradation in performance is most likely attributed to the IMU and vehicle type. Finally, since R-AVO depends directly on the IMU, its performance follows the same trend.

It should be noted that increased vehicle vibrations also affect the VO localization estimate, mainly due to the effect of the added noise on feature detection and matching [31, 44].

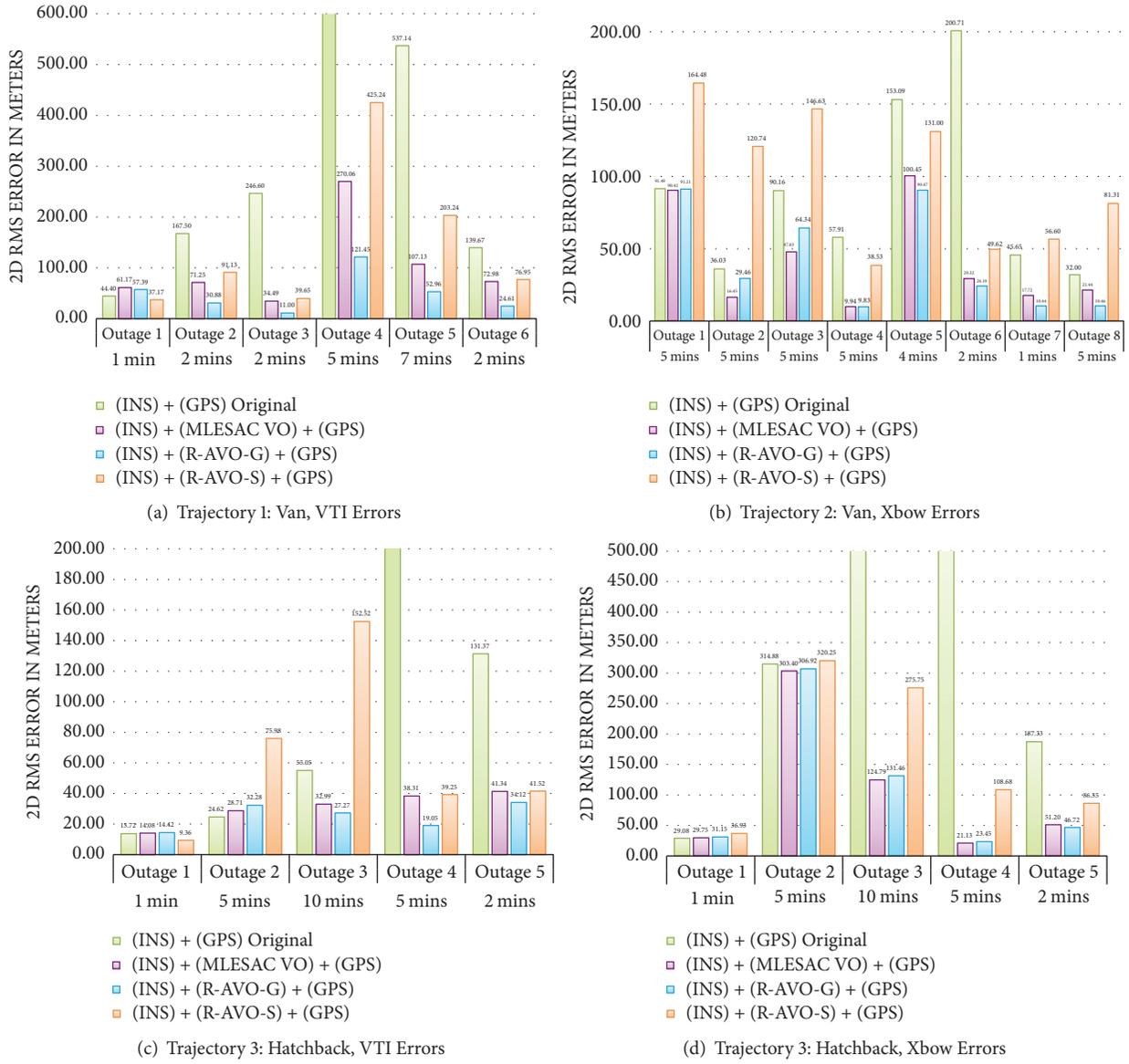


FIGURE 8: Trajectory 2D RMS position errors.

TABLE 2: Comparison of overall trajectory accuracies.

	Van VTI	Van Xbow	Hatchback VTI	Hatchback Xbow
RMS error	1.17 m	2.32 m	3.2 m	3.3 m

A deeper analysis into error correlation, propagation, and compensation between RISS and VO will be investigated as part of future work.

4.5. State Covariance Convergence. GPS provides direct estimates of the vehicle’s position, as seen in the measurement design matrix H_{GPS} of (5). Given such observability, GPS updates cause the latitude and longitude error state covariances to almost instantly converge. This results in good positioning accuracy the moment GPS becomes available. However, GPS does not directly observe the attitude error

states, azimuth in particular. The system needs considerable time with GPS availability for the azimuth error state covariance to converge. A comprehensive treatment of INS/GPS observability can be found in [45].

Since the azimuth error angle is only weakly observable, although the position estimate is accurate during GPS availability, the overall navigation accuracy suffers when several outages occur in succession. This is the highlight of using VO. Since the azimuth angle is directly observable by VO, shown by the design matrix H_{VO} in (6), the system retains convergence even during intermittent GPS availability.

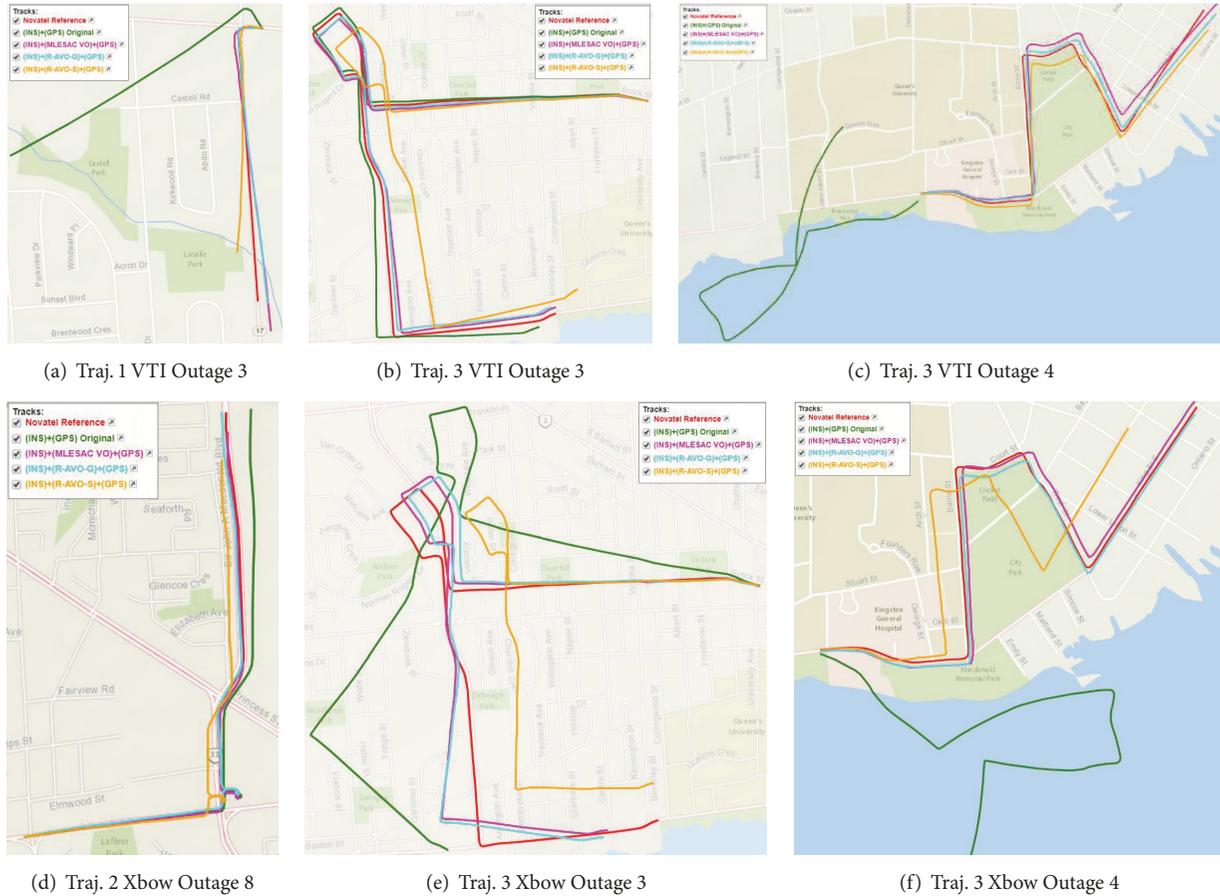


FIGURE 9: Outage portions overlaid on map.

A major consequence of slow azimuth convergence is that the entire INS/GPS navigation solution becomes unusable, highlighted by the large error spikes in the INS/GPS solution across trajectories, shown in Figure 8. In situations where there are short periods of GPS availability followed by considerable outages, the dead-reckoning INS, being embedded within an indirect EKF framework, increments upon an erroneous attitude estimate, which in turn causes significant errors. Given the operating environment of AVs, intermittent GPS availability is common, and, without the added VO azimuth estimate, INS/GPS alone is unsuitable.

4.6. Computational Complexity. Since R-AVO-S uses the Sampson error metric, which is algebraic and straightforward to implement, it has the lowest computation time among the three VO methods. Next comes the R-AVO-G method, which is more computationally intensive than R-AVO-S due to the triangulation involved. Finally, the MLESAC based approach requires several iterations and is more demanding than R-AVO-G.

The average computation time needed for each VO outlier rejection paradigm during the outages is plotted in Figure 10. The spread of computation times is an important parameter to consider, since it indicates whether or not the VO solution

will be used. If a particular epoch requires too much time, a real-time system might discard the delayed solution. Due to the iterative nature of MLESAC, there is large variation in the time required. R-AVO, on the other hand, is always a single iteration, which limits the spread of time needed.

5. Conclusion

The impact of this work lies in presenting a positioning and navigation system for autonomous vehicles that is robust and reliable. The developed INS/VO/GPS integration utilizing reduced-aided VO (R-AVO) was shown to outperform the classic INS/GPS integration algorithm. Particularly for intermittent GPS availability, which is often the case for vehicles in downtown cores, the proposed system had faster error state convergence, which lead to a usable navigation solution even for extended and frequent outages.

The R-AVO algorithm further improved classic VO, showing reduced computational requirements for the outlier rejection stage. The accuracy of R-AVO was similar to the MLESAC based outlier rejection scheme when the vehicle used was a hatchback and the IMU used consisted of very low grade MEMS sensors. In the case of a larger vehicle or a slightly better IMU (although still MEMS) the

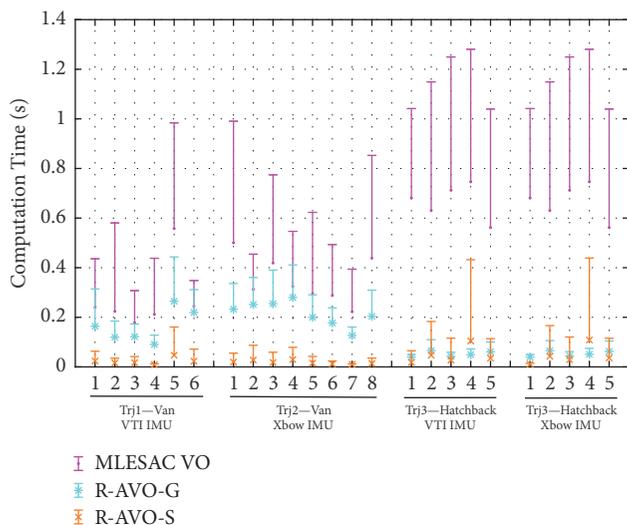


FIGURE 10: A comparison of the mean computation time needed for the VO solutions during each outage. The whiskers represent one standard deviation above the mean processing times.

R-AVO based method outperformed the MLESAC based one.

Future work would include statistical methods to judge the error threshold in the R-AVO based scheme, making use of INS error modelling techniques. Furthermore, the R-AVO method can be extended to include different systems, improving the outlier rejection capabilities of other feature based algorithms that use LiDAR or Radar data.

Data Availability

The data is available upon request. Kindly refer to the article's first page for the authors' contact details.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was funded by Natural Science and Engineering Council of Canada as a discovery grant to Dr. Noureldin.

References

- [1] T A Press, *World's first self-driving taxis start taking passengers*, news, CBC, 2016.
- [2] D. J. Fagnant and K. Kockelman, "Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations," *Transportation Research Part A: Policy and Practice*, vol. 77, pp. 167–181, 2015.
- [3] G. Xu, *GPS: Theory, Algorithms, and Applications*, Springer, Berlin, Germany, 2nd edition, 2007.
- [4] M. M. E. Moussa, A. Osman, M. Tamazin, M. Korenberg, and A. Noureldin, "Enhanced GPS narrowband jamming detection using high-resolution spectral estimation," *GPS Solutions*, vol. 21, no. 2, pp. 475–485, 2017.
- [5] M. Cossaboom, J. Georgy, T. Karamat, and A. Noureldin, "Augmented Kalman filter and map matching for 3D RISS/GPS integration for land vehicles," *International Journal of Navigation and Observation*, vol. 2012, Article ID 576807, 16 pages, 2012.
- [6] M. M. Atia, T. Karamat, and A. Noureldin, "An enhanced 3D multi-sensor integrated navigation system for land-vehicles," *Journal of Navigation*, vol. 67, no. 4, pp. 651–671, 2014.
- [7] A. Noureldin, T. B. Karamat, M. D. Eberts, and A. El-Shafie, "Performance enhancement of MEMS-based INS/GPS integration for low-cost navigation applications," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 3, pp. 1077–1096, 2009.
- [8] D. Nister, O. Naroditsky, and J. Bergen, "Visual odometry," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, pp. 652–659, Washington, DC, USA.
- [9] M. M. Atia, S. Liu, H. Nematallah, T. B. Karamat, and A. Noureldin, "Integrated indoor navigation system for ground vehicles with automatic 3-D alignment and position initialization," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 4, pp. 1279–1292, 2015.
- [10] B. Aleksandr, J. F. Gardner, and P. State, "Constrained navigation algorithms for strapdown inertial navigation systems with reduced set of sensors," in *Proceedings of the 1998 American Control Conference, ACC 1998*, pp. 1848–1852, USA, June 1998.
- [11] U. Iqbal, A. F. Okou, and A. Noureldin, "An integrated reduced inertial sensor system - RISS / GPS for land vehicle," in *Proceedings of the 2008 IEEE/ION Position, Location and Navigation Symposium, PLANS*, pp. 1014–1021, USA, May 2008.
- [12] A. Noureldin, T. B. Karamat, and J. Georgy, "Fundamentals of inertial navigation, satellite-based positioning and their integration," *Fundamentals of Inertial Navigation, Satellite-Based Positioning and their Integration*, pp. 1–313, 2013.
- [13] J. Georgy, A. Noureldin, M. J. Korenberg, and M. M. Bayoumi, "Low-cost three-dimensional navigation solution for RISS/GPS integration using mixture particle filter," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 2, pp. 599–615, 2010.
- [14] S. Gauglitz, T. Höllerer, and M. Turk, "Evaluation of interest point detectors and feature descriptors for visual tracking," *International Journal of Computer Vision*, vol. 94, no. 3, pp. 335–360, 2011.
- [15] K. MacTavish, M. Paton, and T. D. Barfoot, "Night Rider: Visual Odometry Using Headlights," in *Proceedings of the 2017 14th Conference on Computer and Robot Vision (CRV)*, pp. 314–320, Edmonton, AB, May 2017.
- [16] L. Lorigo, R. Brooks, and W. Grimsou, "Visually-guided obstacle avoidance in unstructured environments," in *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robot and Systems. Innovative Robotics for Real-World Applications. IROS '97*, pp. 373–379, Grenoble, France.
- [17] Y. Wang, S. Liu, C. Chen, and B. Zeng, "A hierarchical approach for rain or snow removing in a single color image," *IEEE Transactions on Image Processing*, vol. 26, no. 8, pp. 3936–3950, 2017.
- [18] M. J. Milford and G. F. Wyeth, "SeqSLAM: Visual route-based navigation for sunny summer days and stormy winter nights," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1643–1649, 2012.
- [19] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, UK, 2nd edition, 2004.

- [20] H. C. Longuet-higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature*, vol. 293, no. 5828, pp. 133–135, 1981.
- [21] D. Nistér, "An efficient solution to the five-point relative pose problem," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 756–770, 2004.
- [22] D. Ortín and J. M. M. Montiel, "Indoor robot motion based on monocular images," *Robotica*, vol. 19, no. 3, pp. 331–342, 2001.
- [23] D. Scaramuzza, "1-point-RANSAC structure from motion for vehicle-mounted cameras by exploiting non-holonomic constraints," *International Journal of Computer Vision*, vol. 95, no. 1, pp. 74–85, 2011.
- [24] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [25] P. H. S. Torr and A. Zisserman, "MLESAC: a new robust estimator with application to estimating image geometry," *Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 138–156, 2000.
- [26] P. J. Huber, "Projection pursuit," *The Annals of Statistics*, vol. 13, no. 2, pp. 435–525, 1985.
- [27] B. J. Tordoff and D. W. Murray, "Guided-MLESAC: faster image transform estimation by using matching priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1523–1535, 2005.
- [28] C.-M. Cheng and S.-H. Lai, "A consensus sampling technique for fast and robust model fitting," *Pattern Recognition*, vol. 42, no. 7, pp. 1318–1329, 2009.
- [29] S. Jia, K. Wang, X. Li, and T. Xu, "A novel improved probability-guided RANSAC algorithm for robot 3D map building," *Journal of Sensors*, vol. 2016, Article ID 3243842, 18 pages, 2016.
- [30] S. I. Roumeliotis, A. E. Johnson, and J. F. Montgomery, "Augmenting inertial navigation with image-based motion estimation," in *Proceedings of the IEEE International Conference on Robotics & Automation*, pp. 4326–4333, Washington, DC, USA, May 2002.
- [31] K. Konolige, M. Agrawal, and J. Solà, "Large-scale visual odometry for rough terrain," *Springer Tracts in Advanced Robotics*, vol. 66, pp. 201–212, 2010.
- [32] J.-P. Tardif, M. George, M. Laverne, A. Kelly, and A. Stentz, "A new approach to vision-aided inertial navigation," in *Proceedings of the 23rd IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS 2010*, pp. 4161–4168, Taiwan, October 2010.
- [33] S. Weiss and R. Siegwart, "Real-time metric state estimation for modular vision-inertial systems," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '11)*, pp. 4531–4537, Shanghai, China, May 2011.
- [34] M. J. Veth, R. K. Martin, and M. Pachter, "Anti-temporal-aliasing constraints for image-based feature tracking applications with and without inertial aiding," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 8, pp. 3744–3756, 2010.
- [35] J. Ma, S. Susca, M. Bajracharya, L. Matthies, M. Malchano, and D. Wooden, "Robust multi-sensor, day/night 6-DOF pose estimation for a dynamic legged vehicle in GPS-denied environments," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 619–626, Saint Paul, Minn, USA, 2012.
- [36] T. B. Karamat, R. G. Lins, S. N. Givigi, and A. Noureldin, "Novel EKF-Based Vision/Inertial System Integration for Improved Navigation," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 1, pp. 116–125, 2018.
- [37] M. Maimone, Y. Cheng, and L. Matthies, "Two years of visual odometry on the Mars Exploration Rovers," *Journal of Field Robotics*, vol. 24, no. 3, pp. 169–186, 2007.
- [38] P. Corke, J. Lobo, and J. Dias, "An introduction to inertial and visual sensing," *International Journal of Robotics Research*, vol. 26, no. 6, pp. 519–535, 2007.
- [39] Z. F. Syed, P. Aggarwal, X. Niu, and N. El-Sheimy, "Civilian vehicle navigation: Required alignment of the inertial sensors for acceptable navigation accuracies," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 6, pp. 3402–3412, 2008.
- [40] T. B. Karamat, *Improved Land Vehicle Navigation and GPS Integer Ambiguity Resolution Using Enhanced Reduced-IMU/GPS Integration [Ph.D. thesis]*, Queens University, 2014.
- [41] Y. Sun, M. T. Rahman, T. B. Karamat, A. Noureldin, and Y. Gao, "Integrating vision based navigation with INS and GPS for land vehicle navigation in challenging environments," in *Proceedings of the 29th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS 2016*, pp. 1312–1321, usa, September 2016.
- [42] A. Schneider, *Gps visualizer*, Online, 2016, <http://www.gpsvisualizer.com>.
- [43] A. S. El-Wakeel, J. Li, M. T. Rahman, A. Noureldin, and H. S. Hassanein, "Monitoring road surface anomalies towards dynamic road mapping for future smart cities," in *Proceedings of the 2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 828–832, Montreal, QC, Canada, November 2017.
- [44] D. Scaramuzza and R. Siegwart, "Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles," *IEEE Transactions on Robotics*, vol. 24, no. 5, pp. 1015–1026, 2008.
- [45] S. Hong, M. H. Lee, H.-H. Chun, S.-H. Kwon, and J. L. Speyer, "Observability of error states in GPS/INS integration," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 2, pp. 731–743, 2005.

