Asynchronous vehicle pose correction using visual detection of ground features

Randy E S Harnarinesingh and Chanan S Syan
Department of Mechanical and Manufacturing Engineering
The University of the West Indies
St. Augustine, Trinidad
E-mail: chanan.syan@sta.uwi.edu

Abstract. The inherent noise associated with odometry manifests itself as errors in localization for autonomous vehicles. Visual odometry has been previously used in order to supplement classical vehicle odometry. However, visual odometry is limited in its ability to reduce errors in localization for large travel distances that entail the cumulative summing of individual frame-to-frame image errors. In this paper, a novel machine vision approach for tiled surfaces is proposed to address this problem. Tile edges in a laboratory environment are used to define a travel trajectory for the Quansar Qbot (autonomous vehicle) built on the iRobot iRoomba platform with a forward facing camera. Tile intersections are used to enable asynchronous error recovery for vehicle position and orientation. The proposed approach employs real-time image classification and is feasible for error mitigation for large travel distances.

The average position error for an 8m travel distance using classical odometry was measured to be 0.28m. However, implementation of the proposed approach resulted in an error of 0.028m. The proposed approach therefore significantly reduces pose estimation error and could be used to supplement existing modalities such as GPS and Laser-based range sensors.

1. Introduction
Mobile robots are utilized for a wide range of applications including interplanetary exploration [1], bomb disposal [2] and motorized wheelchair platforms for disabled individuals [3]. The success of a mobile robot in a particular task however requires the robots to have accurate knowledge of its position and orientation (pose). This is referred to as localization and is therefore a fundamental requirement for mobile robots. Localization can be realized using a host of technologies and approaches. The common approaches to mobile robot localization are discussed further.

Shaft encoders are electromechanical devices that can be used to measure angular position and velocity of rotating shafts. Multiple shaft encoders can be utilized to provide estimates of travel velocity and heading for mobile robots which can be used to iteratively update the position of a mobile robot with reference to a previous position. This is termed relative localization.

Shaft encoder readings however are typically noisy and contribute to errors in the determination of robot pose. These errors are cumulative and become significant for large travel distances [4]. Kalman filtering and particle filtering have been applied to treat with encoder based noise for robot localization problems [5]. These approaches however only mitigate errors and cannot eliminate them. The implementation of algorithms such as Kalman filtering also increases the computational complexity of the localization program.

1 To whom any correspondence should be addressed.
In addition to sensor noise, encoder-based localization is hampered by wheel slippage. Slippage is caused by momentary reductions in wheel to surface traction. In the event of slippage, the wheel spins but vehicular motion is not realized. The rotation of the wheel however is measured by the encoders and contributes to pose update. Slippage therefore makes it difficult for a localization systems based on shaft encoders to perform reliably on low traction surfaces such as sand or gravel paths. Sonar, infrared and laser-based technologies are however robust to wheel slippage and have been applied to the problem of mobile robot localization.

Sonar sensors are used in [6] to determine the pose of a mobile robot. The implemented algorithm detected straight segments from the obtained ranges and compared it to a known reference model of the room. The algorithm was able to discriminate between straight line room segments such as walls and smaller elements such as furniture and persons. This allowed for model building to exclude furniture such as tables and desks.

A network of coded infrared beacons are used in [4] whose coordinates were known a priori to the robot. Infrared signals were detected using an upward facing CCD camera and an infrared receiver. Infrared signals are visible to digital CCD cameras due to the phenomena of baseband replication that arises from signal sampling. Detection of a coded infrared transmission identified the position of the transmitter and allowed for triangulation-based absolute localization of the vehicle.

Laser-based Range Finding (LRF), also known as Laser Radar (LIDAR) is the process of illuminating an object with laser and using the reflected signal to determine distance. LRF allows for high accuracy range detection and has been applied to the problem of mobile robot localization [7]. LRF approaches are advantageous because they do not necessarily require the addition of artificial beacons to an environment like those used in [4]. This allows for rapid deployment of robots in unknown surroundings without the need for environmental modifications.

The high accuracy of LRF allows makes it possible to localize a robot without knowledge of the robot kinematic model [8]. Successful localization without the need for robot kinematic modelling reduces the conceptual and computational complexity of odometry. LRF also permits the localization of mobile robots without prior knowledge of an environmental map [9]. This is useful in disaster recovery scenarios. For example, earthquake damaged buildings do not lend themselves to rapid mapping prior to ad hoc exploratory and recovery missions.

The inherent advantages of LRF are however outweighed by the costly nature of laser transmitters and receivers. This renders LRF inappropriate for low cost mobile platforms. LRF also entails high computational expense when map building algorithms are implemented. In addition, active sensing technologies such as laser, sonar and infrared possess common intrinsic disadvantages. Robots can misinterpret a signal transmitted from another device as a reflected signal originating from itself. This can corrupt the localization ability of active sensors. Passive sensors such as cameras address this problem. Cameras are an advantageous sensor for localization because they provide a large amount of information and are not susceptible to the noise inherent to light and sound based sensors [10].

Cameras are used in [11] to implement a Simultaneous Localization and Mapping (SLAM) algorithm using naturally occurring scale-invariant landmarks in an environment. The invariance of environmental landmarks to image translation, scaling and rotation enhanced their suitability for mobile robot localization and map building since it allowed for feature construction from various viewpoints. However, the algorithm required 2-3 seconds for the processing of acquired stereo images. Localization would therefore be achieved for stationary robots however, a travelling robot would realize a localization error because of the distance covered during the attendant image processing.

The use of cameras for localization in general requires the use of image databases [12], filtering algorithms [13] or geometric image models which increase the computational expense of the localization system. A hybrid approach to mobile robot localization using shaft encoders and a monocular camera is proposed in this paper. The hybrid approach averts the need for image databases, geometric image models and filtering algorithms and relies on the inherent visual features of a tiled laboratory environment to supplement encoder-based localization.
2. Hybrid approach to robot localization
The hybrid approach to mobile robot localization is illustrated in Figure 2. The main components of this approach are described further.

2.1. Colored line following
The colored line following approach is designed for tiled environments in which the tile edges are color-wise distinguishable to the inner tile portions. Colored line following is implemented to follow an externally defined trajectory for the mobile robot and to allow for the acquisition of intersections for localization. This requires the use of a forward facing monocular camera. The Quanser Qbot is a mobile differential drive robot built on the iRobot iRoomba platform. The Qbot features a forward facing monocular camera and is used for this study. The approach however may be modified to operate with ceiling tiles that satisfy the color-wise discrimination requirement using an upward facing monocular camera. In general, any floor visual features that are repetitive and overlay the floor surface can be utilized. The tiled environment used for testing the hybrid approach is pictured in Figure 1 using a smartphone camera.

Figure 1. Tiled Laboratory surface pictured using a smartphone camera
2.1.1. Region of Interest Extraction. The black line as viewed from the forward facing camera on the Qbot is pictured in Figure 3. The entire image however does not have to be utilised for the coloured line following approach. The important image data resides at the bottom centre of the image where the line is immediately in front of the robot. This area is outlined in red in Figure 3. In this regard, a Region-of-Interest-Extraction (ROIE) is implemented to acquire this image segment. This has the benefit of reducing the computational expense of the later processing stages.

![Diagram of mobile robot localization](image)

Figure 2. Hybrid approach to mobile robot localization

Wheel velocities: \((V_{\text{LEFT}}, V_{\text{RIGHT}})\)
Figure 3. Black line tile edge pictured from forward facing Qbot camera with outline Region of Interest

The ROIE image is illustrated in Figure 4.

Figure 4. Region of Interest Extracted Image of Black line tile edge

2.1.2. Thresholding. Thresholding was implemented in order to discriminate the black line tile edges from the surrounding tile interior. The mathematical description of thresholding is given in (1).

\[
\begin{align*}
    h(x, y) = \begin{cases}
        255, & \text{if } th1 \leq f(x, y) \leq th1 \\
        0, & \text{otherwise}
    \end{cases}
\end{align*}
\]

(1)

The thresholding ranges were determined experimentally using grayscale histograms. The thresholded image is presented in Figure 5.

Figure 5. Thresholded image of black line tile edge

2.1.3. Dilation. Discontinuities in the thresholded image are possible due to glare and dirt on the tile edges. This is responsible for the variable width of the detected line in Figure 5. This presents
a challenge to intersection detection which is the basis for the localization approach. Dilation is a morphological operation that can be used to remove discontinuities in images due to noise. The mathematical description of dilation is presented in (2).

\[ A, B \in \mathbb{Z}^2 \]
\[ \text{dilation}(A, B) = \{ z | (B)_z \cap A \neq \emptyset \} \text{ where} \]
\[ \hat{B} = \{ w | w = -b, f \text{ or } b \in B \}, \text{ and} \]
\[ (\hat{B})_z = \{ c | c = w + z, f \text{ or } w \in \hat{B} \} \]

(2)

Figure 6 and Figure 7 present the case of thresholded images of the same tile intersection before and after dilation respectively.

![Image](image1.png)

**Figure 6.** Thresholded image before dilation operation

![Image](image2.png)

**Figure 7.** Thresholded image after dilation

A blob is define as a contiguous set of pixels of the same color. Analysis was performed on the dilated image in order to determine the centroid and width of the largest blob. The centroid of the largest pixel was mapped to the spatial domain and input to the Inverse Kinematic model of the vehicle in order to obtain the wheel velocities for motion. The width of the blob was used to detect intersections for localization. The derivation of the robot kinematic model is presented in Section 3.

3. Differential Drive Kinematic Modelling

The Quanser Qbot is a differential drive mobile vehicle consisting of two wheels mounted on a common axis. Differential drive vehicles possess zero-minimum turning radius which is advantageous since it allows rotation in tight spaces. Each wheel can be controlled independently.
in order to navigate the vehicle. However, the vehicle must rotate about a point that lies along the common left and right wheel axes. This point is known as the Instantaneous Centre of Curvature (ICC) and is depicted in Figure 8 (Dedek and Jenkin, 2000).

\[ \text{ICC} \]

\[ \omega \]

\[ R \]

\[ d/2 \]

\[ d/2 \]

\[ (x,y) \]

\[ V = R \omega \]

\[ V_L \]

\[ V_R \]

\[ x \]

\[ y \]

\[ x \]

\[ y \]

\[ \theta \]

\[ (x,y) \]

\[ d/2 \]

\[ d/2 \]

\[ V \]

\[ R \]

\[ \omega \]

\[ V_L \]

\[ V_R \]

Figure 8. Differential drive kinematic model of Quanser Qbot

The rate of angular rotation about the ICC is the same for both wheels. The following equations therefore establish a relationship between the parameters of motion for a differential drive vehicle:

\[ V_R = \omega (R + \frac{d}{2}) \tag{3} \]

\[ V_L = \omega (R - \frac{d}{2}) \tag{4} \]

Equations (3) and (4) are independent equations in \( R \) and \( \omega \), the Radius of Curvature (ROC) and the angular velocity, and can therefore be uniquely solved to give the following expressions:

\[ \omega = \frac{(V_R - V_L)}{d} \tag{5} \]

\[ R = \frac{d}{2} \times \frac{(V_R + V_L)}{(V_R - V_L)} \tag{6} \]

Equations (5) and (6) can be used to determine the instantaneous tangential velocity \( V \) of the vehicle.

\[ V(t) = R(t)\omega(t) = \frac{d}{2} \times \left[ \frac{V_R(t) + V_L(t)}{V_R(t) - V_L(t)} \right] \times \left[ \frac{[V_R(t) - V_L(t)]}{d} \right] = \frac{[V_R(t) + V_L(t)]}{2} \tag{7} \]
3.1.1. Forward Kinematics. The forward kinematics problem deals with finding the vehicle position and orientation \((x, y, \theta)\) at time \(t\) using the initial position \((x_0, y_0, \theta_0)\) and the wheel velocities. The forward kinematic solution in discrete-time form is presented in (8) for the case of encoder-based odometry only. Discrete-time form is necessary since the Qbot Application Programming Interfaces (APIs) for accessing the camera output and assigning the wheel velocities are MATLAB/Simulink.

\[
\begin{bmatrix}
  x[k+1] \\
  y[k+1] \\
  \theta[k+1]
\end{bmatrix} =
\begin{bmatrix}
  x[k] + \frac{1}{2} (V_R + V_L) \cos(\theta) T_S \\
  y[k] + \frac{1}{2} (V_R + V_L) \sin(\theta) T_S \\
  \theta[k] + \frac{1}{d} (V_R - V_L) T_S
\end{bmatrix}
\]  

(8)

The hybrid approach however relies on the detection of tile edge intersections for localization. Tile edge intersection are equally spaced at 1.23m. The approach constrains the robot to start at an intersection facing the black tile edge. The starting intersection is taken as the origin regardless of which intersection is used. The Forward Kinematic update rule for intersection detection is given in (9).

\[
\begin{bmatrix}
  x[k+1] \\
  y[k+1] \\
  \theta[k+1]
\end{bmatrix} =
\begin{bmatrix}
  x_{\text{previous intersection}} + 1.23 \\
  0 \\
  0
\end{bmatrix}
\]  

(9)

3.1.2. Inverse Kinematics. The inverse kinematics problem deals with finding the required wheel velocities to obtain a specified vehicle pose \((x, y, \theta)\) at time \(t\). This problem is modified in the discrete-time case as finding the required wheel velocities to get the vehicle to pose \((x_{k+1}, y_{k+1}, \theta_{k+1})\) from \((x_k, y_k, \theta_k)\) where \(k\) is the discrete-time index.

The vehicle pose at discrete-time indices \(k\) and \(k+1\) allows for the determination of the distance and the angular heading \(\alpha\) between the two poses.

\[
distance = \sqrt{(x_{k+1} - x_k)^2 + (y_{k+1} - y_k)^2} 
\]

(10)

\[
\alpha = \theta_{k+1} - \theta_k 
\]

(11)

The distance and heading between the two poses can be used to determine \(V\) and \(\omega\).

\[
V = \frac{\text{distance}}{T_S} 
\]

(12)

\[
\omega = \frac{\alpha}{T_S} 
\]

(13)

The left and right wheel velocities are then obtained from \(V\) and \(\omega\).

\[
V_R = \frac{2V + d \omega}{2} 
\]

(14)

\[
V_L = 2V - V_R 
\]

(15)
In this approach, \((x_k, y_k, \theta_k)\) is assumed to be \((0,0,0)\) and \((x_{k+1}, y_{k+1}, \theta_{k+1})\) are the coordinates of the largest blob in the dilated image mapped to the spatial domain.

4. Results

4.1. Colored line following

The coloured line following approach was used to drive the vehicle for 100 executions of 8m travel distances. The vehicle remained on the black line without deviation for all executions.

4.2. Hybrid approach to mobile robot localization

The vehicle was driven for 30 executions of 8m of black line following. Figure 9 presents the average trajectory plot for forward kinematics using shaft encoders only. Figure 10 present the average trajectory plots for forward kinematics using the hybrid approach.

![Figure 9. Average Qbot trajectory based solely on shaft encoders](image-url)
5. Discussion
Colored line following was implemented for following the black line tile edges. The colored line following approach worked reliably for 100 executions. This validated the development of the colored line following subsystems.

Figure 9 presents the average trajectory for the Qbot using encoder-based Forward Kinematics. Localization drift is evident in the trajectory plot and grows as the travel distance is increased as suggested in [4]. The absolute position error at the end of travel was 0.28m.

Figure 10 presents the average trajectory using the hybrid approach to localization. Discontinuities are evident in these trajectory plots. These are due to the auto-corrective feature of the approach. For example, if the vehicle has travelled the length of 1 intersection which is 1.23m but encoder-based odometry measures a distance not equal to 1.23m, the position correction feature is manifested as a finite jump in the trajectory plot or a discontinuity. This asynchronous adjustment mechanism ensures that the error for any possible travel distance is arrested to a finite band. The absolute error at the end of travel was 0.028m.

The proposed approach therefore resulted in a 10-fold reduction in position error over the traversed space. This is a significant reduction in error and validates the proposed approach for tiled environments.

6. Conclusion and future work
The proposed hybrid approach to localization resulted in a ten-fold reduction in position error from 0.28m to 0.028m by supplementing the encoder-based forward kinematic system with the visual detection of ground tile intersections. This ensures that the autonomous vehicle has an accurate report of its position which is essential for successful task completion. The proposed approach is therefore feasible and could be adopted to supplement existing modalities such as GPS and LRF.

Future work will investigate the feasibility of grid based navigation using the Qbot. In the current implementation, the Qbot detects intersection approximately 450mm ahead of travel but travels past the intersection. Grid based navigation would require the vehicle to stop at
intersections and allow turning to face other intersections before resuming travel. This presents a challenge due to the error in the measurement of rotation angle by shaft encoders. However the results of this paper motivate a strategy for intersection turning based solely on visual feedback.

References

[1] Volpe R, Balaram J, Ohm T and Ivlev R 1997 The Rocky 7 Mars Rover Prototype Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (Osaka, Japan, 4-8 November 1996)

[2] Costo S and Molfino R 2004 A New Robotic Unit Ror Onboard Airplanes Bomb Disposal Proc. 35th International Symposium on Robotics (Paris, France, 23-26 March 2004)

[3] Iturrate I, Antelis J, Kubler A and Minguez J 2009 A Non-Invasive Brain-Actuated Wheelchair Based on a P300 Neurophysiological Protocol and Automated Navigation IEEE Transactions on Robotics 25(3) 614-627

[4] Brassart E, Pegard C and Mouaddib M 2000 Localization using infrared beacons Robotica 18 153-161

[5] Von Der Hardt H, Arnould P, Wolf D and Dufaut M 1994 A method of mobile robot localisation by fusion of odometric and magnetometric data The International Journal of Advanced Manufacturing Technology 9(1) 65-69

[6] Drumheller M 1987 Mobile Robot Localisation Using Sonar IEEE Transactions on Pattern Analysis and Machine Intelligence 9 325-332

[7] Ribeiro M I and Goncalves 1996 Natural landmark based localisation of mobile robots using laser range data Proc. of the First Euromicro Workshop on Advanced Mobile Robot (Kaiserslautern, Germany, 9-11 Oct 1996) pp 186-193

[8] Spero D J and Jarvis R A 2004 Toward Exteroceptive Based Localisation Proc. of the 2004 IEEE International Conference on Robotics, Automation and Mechatronics (Singapore, 1-3 December 2004) pp 822-827

[9] Zhang D, Kurazume R, Iwashita Y and Hasegawa T 2011 Robot localization under perceptual aliasing conditions based on laser reflectivity using particle filter Proc of the 2011 IEEE/SICE International Symposium on System Integration (Kyoto, Japan, 20-22 December 2011) pp 90-95

[10] Wolf J, Burgard W and Burkhardt H 2002 Using an Image Retrieval System for Vision-Based Mobile Robot Localization Lecture Notes in Computer Science 2383 108-119

[11] Se S, Lowe D and Little J 2002 Mobile Robot Localization and Mapping with Uncertainty using Scale-Invariant Visual Landmarks International Journal of Robotics Research 21 735-758

[12] Lamon P, Nourbaksh I, Jensen B and Siegwart R 2001 Deriving and matching image fingerprint sequences for Mobile Robot Localization Proc. of the 2001 IEEE International Conference on Robotics and Automation 2 pp 1609-1614

[13] Georgiev A and Allen P K 2004 Localization methods for a mobile robot in urban environment IEEE Transactions on Robotics 20(5) 851-864