Distance Estimation of Monocular Based on Vehicle Pose Information

S H Qi¹, J Li¹, Z P Sun¹, J T Zhang¹ and Y Sun¹
¹National University of Defense Technology, Changsha, China

Abstract. In order to measure the distance between our vehicle and the target vehicle by monocular vision and eliminate the estimation error bring by changing of vehicle pose, we propose the distance estimation method based on the vehicle pose information, which can be used to eliminate the error of distance estimation effectively cause by the change of the pitch angle and roll angle of unmanned vehicle. In addition, the pose information could also help us estimate whether the vehicle is in a slope, thus engage in distance estimation for the vehicles. Several groups of data was collected for the experiments. And the result prove the validity of the algorithm in distance estimation.

1. Introduction

With the development of unmanned vehicle, the intelligence of vehicles has been continuously improved. Autopilot via machine vision has gradually become one of the research directions of research teams and companies. Obtaining the external environment information through the camera is one of the important ways to achieve autopilot. Meanwhile, using the monocular camera to measure the distance between our vehicle and the front vehicle in real time can not only assist the driving, but also provide favorable information for path planning and driving safety of unmanned vehicles.

Monocular vision distance estimation utilizes image captured by a single camera to compute distance. Currently, it is divided into model-based distance estimation method and depth estimation based on deep learning. In the traditional distance estimation method based on model, the pinhole imaging model of the camera is mainly used. Stein G P et al. [1] and Gat I et al. [2] detected the position of the vehicle in the image by traditional vehicle detection algorithm firstly, and then carries out distance estimation of the currently detected vehicle according to the principle of pinhole camera. However, the accuracy of the position is particularly important for distance estimation since the position of the vehicle bottom should be known. Meanwhile, the position of vanish line in the image is also necessary. However, the bump during running is unavoidable which will cause error in distance estimation. Dagan E et al. [3] proposed a method to calculate the current vehicle distance according to the size of width of the rear of vehicle in the image. But the width of the vehicle should be known in advance should this method be adopted. Based on the position and width information of vehicle, Kim G et al. [4] proposed a distance estimation algorithm in view of the combination of the bottom position of vehicle and vehicle width information. Guo L et al. [5] calculated the longitudinal distance to preceding vehicles according to the geometric mapping relationship between the image coordinate system and the world coordinate system. Shen Z X et al. [6] established a mathematical regression model utilizing the relationship between the distance and the height of the target in the image, then quadratic curve was used to measure the distance from the regression model. Chen Y et al. [7] raised a model with greater suitability to solve distance estimation of straight line and curve. However, the above traditional distance estimation methods have not take the changing of extrinsic parameter of...
camera into consideration which was caused by thrashing during driving and would affect the accuracy of the distance estimation. In the distance estimation algorithm based on deep learning, such as DORN [8] and CSWS_E_ROB [9], a general network model was proposed to predict the depth estimation of the entire picture. Although the depth estimation is better than previous methods, the calculation on target distance estimation of unmanned vehicle is redundant and untargeted. In addition, it takes 0.5 second and 0.2 second respectively to finish the expensive computation which makes it impossible to apply these methods in vehicle range estimation simultaneously.

In order to solve the distance estimation error caused by the change of vehicle pose during the driving, therefore we take the method in [1, 2] as baseline and propose a distance estimation algorithm based on the vehicle pose information. The Inertial Measurement Unit (IMU) was adopted to measure the pitch angle and roll angle during driving. Then the angle information would be added in distance estimation to reduce the inaccuracy caused by vehicle shaking. At the same time, it is possible that the vehicle may drive on a slope in real road condition, so we could decide whether introduce IMU information to estimate vehicle distance or not after the analysis of vehicle posture by the pitch angle.

2. Distance Estimation Algorithm Based on Vehicle Pose Information

According to the pinhole imaging model, the monocular vision system can be simplified as a camera projection model [1, 2], as shown in figure 1.

![Figure 1. Schematic diagram of the imaging geometry.](image)

Figure 1 shows a diagram of a schematic pinhole camera. It is consisted of a camera $C$ with focal $f$ and an imaging plane $I$. The height of camera is $H$. The target vehicle is at a distance $Z$ from the camera. The bottom of the target vehicle projects on the image $I$ at a position $y$. $dc$ is the distance between camera and the origin of vehicle coordinate system, $dv$ denotes the distance that we care. The focal and the image coordinates on imaging plane are usually in $mm$. The distance $Z$, $dc$, $dv$ and $H$ is typically in $m$.

2.1 Distance estimation based on the position of vehicle

According to the pinhole imaging model we can estimate distance by the equation:

$$y = \frac{f \ast H}{Z} \quad (1)$$

To compute distance $Z$, it is necessary to turn $y$ on the imaging plane $I$ into the pixel coordinate system of image. The equation can be presented as follows:

$$\begin{align*}
    u &= \frac{X}{dX} + u_0 \\
    v &= \frac{Y}{dY} + v_0
\end{align*} \quad (2)$$

where $(u,v)$ is the coordinate of target point in pixel coordinate system, $(X,Y)$ is the coordinate of the image coordinate system in $mm$, $dX$ and $dY$ represent the physical size of the pixel along the X-axis and Y-axis respectively. $(u_0, v_0)$ represents the coordinate of optical center in pixel coordinate system.
Take equation (2) into equation (1) and combine to figure 1, we can get the following equation:

\[ d_v = \frac{ay \cdot H}{v - v_0} + d_e \]  

(3)

where \( ay = f/dY \) is the scale factor along the vertical axis of the image, it can be acquired from the internal parameter matrix of camera, \( v \) denotes the bottom of target vehicle in the image. \( v_0 \) is the position of vanish line in the image that the point at infinity projects on.

The position of vanish line \( v_0 \) must be known in order to engage in distance estimation. The corresponding relationship between vanish line and internal and external parameters of camera can be deducted from principle of camera perspective as follows:

\[ P_{3 \times 4} = K_{3 \times 3} \cdot [R_{3 \times 3}, t_{3 \times 1}] \]  

(4)

\[
\begin{bmatrix}
    u \\
    v \\
    1
\end{bmatrix} =
\begin{bmatrix}
    P_{11} & P_{12} & P_{13} & P_{14} \\
    P_{21} & P_{22} & P_{23} & P_{24} \\
    P_{31} & P_{32} & P_{33} & P_{34}
\end{bmatrix}
\begin{bmatrix}
    x \\
    y \\
    z \\
    1
\end{bmatrix}
\]  

(5)

According to limit theorem, it can be simplified as follows:

\[
\begin{align*}
    v &= \lim_{y \to \infty} \left( P_{21} \cdot x + P_{22} \cdot y + P_{24} \right) \\
    &= \lim_{y \to \infty} \left( \frac{P_{21}}{P_{32}} \cdot x + \frac{P_{22}}{P_{32}} \cdot y + \frac{P_{24}}{P_{32}} \right) \\
    &= \frac{P_{21}}{P_{32}} \cdot x + \frac{P_{22}}{P_{32}} \cdot y + \frac{P_{24}}{P_{32}} \\
    &= \frac{P_{21}}{P_{32}} \cdot x + \frac{P_{22}}{P_{32}} \cdot y + \frac{P_{24}}{P_{32}} \\
\end{align*}
\]  

(6)

where \( R \) is rotate matrix and \( t \) is translate matrix of extrinsic parameter, \( K \) is intrinsic parameter matrix of camera. \((x,y,z,1)\) is homogeneous coordinates of target in the world coordinate system. \((u,v)\) is pixel coordinate of target in the picture and \( s \) is scale factor for this pixel.

2.2 Update rotate matrix \( R \)

The pitch angle and roll angle of the vehicle is changing when vehicle is driving and it leads to the extrinsic parameter of camera changing. The original calibrated rotation matrix \( \bar{R} \) is not suitable for current status of vehicle and it results in vanish line \( v_0 \) changing. As shown in figure 2 and Table 1, the change in the pitch angle of the vehicle causes even larger fluctuation of vanish line and the estimated distance has a large estimation error with the pitch angle changing. From these results, it is confirmed that the change of the pitch angle has a great influence on the distance estimation.
Table 1. Vanish line and distance estimation with different pitch angle.

| Pitch/° | -2   | -1.5 | -1   | -0.5 | 0    | 0.5  | 1.0  | 1.5  | 2.0  | gTruth |
|---------|------|------|------|------|------|------|------|------|------|--------|
| Vanish line | 314.2 | 331.2 | 348.2 | 365.2 | **382.2** | 399.1 | 416.1 | 433.0 | 449.9 | 384    |
| Distance/m | 18.59 | 20.61 | 23.18 | 26.59 | **31.28** | 38.18 | 49.33 | 70.38 | 125.1 | 34.7   |
| Error/m    | 16.11 | 14.09 | 11.52 | 8.11  | **3.42**  | 3.48  | 14.62 | 35.68 | 90.42 | ——     |

In order to eliminate the estimated error caused by vehicle vibration, inertial measurement unit is introduced to measure the pitch angle \( \alpha \) and roll angle \( \beta \) for updating the rotate matrix \( R \), and then we use equation (6) to correct position of vanish line to improve the accuracy of distance estimation. Due to the manner that the rotate matrix \( R \) generated by the pose angle is not uniform, we define that the rotate matrix is generated in the manner of "yaw angle \( \gamma \) → pitch angle \( \alpha \) → roll angle \( \beta \)" and it is always considered that the counterclockwise direction is the positive direction of the yaw angle in the plan view, upward is the positive direction of the pitch angle and the right side is the positive direction of the roll angle when flipping. The change of pose angle and its size are determined by the vehicle itself and is independent of the selected coordinate system [10]. The coordinate system is defined as shown in figure 3.

If the increment corresponding to the yaw angle, the pitch angle, and the roll angle are \( \Delta \gamma \), \( \Delta \alpha \), and \( \Delta \beta \) respectively, the incremental rotate matrix \( \Delta R \) can be computed as follow:

\[
\Delta R = \begin{bmatrix}
\cos \Delta \beta & 0 & -\sin \Delta \beta \\
0 & 1 & 0 \\
\sin \Delta \beta & 0 & \cos \Delta \beta
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 \\
0 & \cos \Delta \alpha & \sin \Delta \alpha \\
0 & -\sin \Delta \alpha & \cos \Delta \alpha
\end{bmatrix}
\begin{bmatrix}
\cos \Delta \gamma & \sin \Delta \gamma & 0 \\
-\sin \Delta \gamma & \cos \Delta \gamma & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

(7)

However, since the yaw angle does not affect the camera's extrinsic parameter matrix during the driving, the incremental rotate matrix \( \Delta R \) can be simplified as follows:

\[
\Delta R = \begin{bmatrix}
\cos \Delta \beta & 0 & -\sin \Delta \beta \\
0 & 1 & 0 \\
\sin \Delta \beta & 0 & \cos \Delta \beta
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 \\
0 & \cos \Delta \alpha & \sin \Delta \alpha \\
0 & -\sin \Delta \alpha & \cos \Delta \alpha
\end{bmatrix}
\]

(8)

Then the update rotate matrix can be get by follow equation:

\[
R_{update} = R \times \Delta R
\]

(9)

Combine equation (9), equation (4), equation (6) and equation (3) can estimate distance in real-time. Algorithm 1 provides a brief outline of our method.

**Algorithm 1 distance estimation of monocular based on vehicle pose information**

1. Determine whether the vehicle is on a slope or not ;
2. Vehicle is not on a slope, use pitch angle $\alpha$ and roll angle $\beta$ to update $R$ and vanish line $v_0$;
3. Vehicle is on a slope, use calibrated intrinsic parameter matrix and extrinsic parameter matrix to compute vanish line $v_0$;
4. Estimate distance of target vehicle using equation (3).

3. Acquisition for ground truth of vehicle distance
In order to verify the performance of our algorithm, it is necessary to obtain the ground truth of the distance for the target vehicle. We use the millimeter wave radar to get the true value of target distance. The millimeter-wave radar is mounted directly in front of the vehicle and is a 76-77 GHz ESR radar with long-range and medium-range scanning capabilities.

After obtaining the bounding box of the target vehicle and the radar data corresponding to current image, we project radar data to the image according to the projection principle and the result is shown in figure 5. Then the matching algorithm in [11] is adopted to select the radar point corresponding to the bounding box. The blue point is the matching result corresponding to these bounding boxes. When we complete the matching of bounding boxes and the radar data, the ground truth of distance need to be computed. Calculation principle of distance is shown in figure 4, the distance $D_m$ can be acquired by radar. $D_r$ is the position of the millimeter-wave radar in the vehicle coordinate system, and $\theta$ is the azimuth of the target vehicle. $D$ indicates distance between the origin of coordinate system and target vehicle and can be obtained by sine theorem. The true value of the distance along the vertical direction can be obtained. The ground truth of distance is shown in figure 5.

4. Experiment

4.1 Experiment data
Images are captured by the camera mounted on the windshield horizontally while radar data is collected by the millimeter wave radar installed vertically shown in figure 4. The size of image is 1824*940*3. Before our experiment, we need to measure height of camera $H$, the distance $D_r$ and the distance $D_c$. Calibration of camera is realized by exploiting the calibrated method of Zhang Zhengyou [12] contained in MATLAB Toolbox. After calibration is done, we can get the intrinsic parameter matrix $K$ and extrinsic parameter matrix $R, t$. Angle information is acquired by inertial measurement unit. The frequency of camera, radar and IMU are 20Hz.

4.2 Result
To verify the performance of our algorithm, we select two groups of images to test, one is from near to far, the other one is from far to near. Bounding boxes are labeled by Ground Truth Labeler in Matlab 2017b. The test result of distance estimation shown in figure 6.

![Figure 6](image)

From figure 6, the trajectory of the blue point is closer to the yellow trajectory than the trajectory of the red point intuitively. To evaluate the effectiveness of our algorithm, the L1 norm is used as an evaluation criterion.

| Group | baseline | ours | improve | Group | baseline | ours | improve |
|-------|----------|------|---------|-------|----------|------|---------|
| L1 norm/m | 2.658 | 2.203 | 16.9% | L1 norm/m | 3.1922 | 1.6397 | 49.63% |
| Variance | 7.914 | 4.224 | —— | Variance | 12.511 | 3.323 | —— |

From figure 6 and table 2, it can be seen that when vehicle pose information is introduced to estimate distance, it can improve the accuracy of distance estimation and our algorithm is more precise and stable than baseline algorithm without vehicle pose information. In these two algorithms, it can be seen that the estimated result of the near vehicle is more accurate than the far vehicle. The reason is shown in (3), because the farther the vehicle, the smaller the value of \( v-v_0 \). When the pixel error of \( v-v_0 \) is same between near vehicle and far vehicle, the distance error will larger for far vehicle. This problem is bring by the defect of the algorithm itself. Therefore, this algorithm is only suitable for distance estimation of nearby vehicles.

In order to determine the optimal distance estimation range of our algorithm, the following supplementary analysis were made. From table 3, it can be seen that when the target vehicle is within 30 meters, our algorithm has the optimal performance. The distance error and variance of distance estimation reduce from 3.73%, 4.76% to 2.37%, 2.18% respectively. We can make a conclusion that our algorithm is suitable for distance estimation when vehicles are within 30 meters.

### 5. Conclusion

In this paper, a distance estimation algorithm based on the vehicle pose information is proposed to improve the accuracy of distance estimation. Firstly, we update the rotate matrix \( R \) by adding the vehicle pose information, then we update the position of vanish line \( v_0 \) to achieve the goal that eliminates the estimation error caused by vehicle flutter. The advantages and disadvantages of the algorithm are analyzed through experimental data and the results prove that this algorithm is suitable for near target vehicle. The optimal application range of the algorithm for vehicle distance estimation is obtained.

![Table 3](image)

| Group 1 distance/m | 10   | 20   | 30   | 40   | 50   | 60   |
|--------------------|------|------|------|------|------|------|
| baseline           | —    | —    | 0.6271 | 1.1199 | 2.1540 | 2.6574 | 2.6433 |
| L1 norm/m          | 3.14% | 3.73% | 5.38% | 5.31% | 4.41% |
| ours               | 0.6077 | 0.7099 | 1.3806 | 1.9097 | 2.1913 |
error/m | 3.04% | 2.37% | 3.45% | 3.82% | 3.65%
---|---|---|---|---|---
Variance | baseline | 0.2940 | 2.0973 | 5.1837 | 8.3066 | 7.9142
| ours | 0.2326 | 0.6894 | 2.3590 | 2.9303 | 4.2241
Group 2 | range/m | 10 | 20 | 30 | 40 | 50 | 60
| baseline | | | | 1.4275 | 2.3954 | 3.1852 | 3.1852
| ours | | | | | | | |
| L1 norm | error/m | | | | | | |
| baseline | | | | | | | |
| ours | | | | | | | |
| Variance | baseline | | | | | | |
| ours | | | | | | | |

Acknowledgement
This work has been supported by the Nation Natural Science Foundation of China under Grant 61473303.

Conferences
[1] Stein G P, Mano O and Shashua A 2003 Vision-based ACC with a single camera: bounds on range and range rate accuracy *Intelligent Vehicles Symposium Proceedings IEEE* pp120-5
[2] Gat I, Benady M and Shashua A 2005 A Monocular Vision Advance Warning System for the Automotive Aftermarket *Sae World Congress & Exhibition*
[3] Dagan E, Mano O and Stein G P 2004 Forward collision warning with a single camera *Intelligent Vehicles Symposium IEEE* pp 37-42
[4] Kim G and Cho J S 2012 Vision-based vehicle detection and inter-vehicle distance estimation *International Conference on Control, Automation and Systems IEEE*
[5] Guo L, You-Chun X U and Ke-Qiang L I 2006 Study on Real-time Distance Detection Based on Monocular Vision Technique *Journal of Image & Graphics* 11(1) pp 74-81
[6] Shen Z X 2007 Monocular vision distance detection algorithm based on data regression modeling *Computer Engineering & Applications* 43(24) pp 15-7
[7] Chen Y and Chen Y 2014 Calculation method of preceding vehicle distance based on monocular vision *Journal of Mechanical & Electrical Engineering*
[8] Li B, Dai Y and He M 2017 Monocular Depth Estimation with Hierarchical Fusion of Dilated CNNs and Soft-Weighted-Sum Inference
[9] Fu H, Gong M, Wang C, Batmanghelich K and Tao D 2018 Deep Ordinal Regression Network for Monocular Depth Estimation. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* pp 2002-11
[10] Yu Rui 2017 High Precision 3D Mapping and Localization Based on Lidar *National University of Defense Technology*
[11] Liu X, Sun Z and He H 2011 On-road vehicle detection fusing radar and vision *IEEE International Conference on Vehicular Electronics and Safety* pp 150-4
[12] Zhang Z 2000 A flexible new technique for camera calibration *IEEE Transactions on pattern analysis and machine intelligence* 22(11) pp 1330-4