Monocular Vision Distance Measurement Method Based on Dynamic Error Compensation

Geng Lei, *Xiao Zhitaotao, Wu Jun, Zhang Fang, Miao Jinghong
School of Electronics and Information Engineering, Tianjin Polytechnic University, Tianjin 300160, China
(genglei@tjpu.edu.cn, *Corresponding Author zhitaotao@yahoo.com, zhenkongwujun@163.com, hhzhafang@126.com, miaojinghong@tjpu.edu.cn)

Abstract

On the basis of a cost-effective embedded system, this paper gives a preceding vehicle distance measuring method with monocular vision techniques for highway driver assistance system. The road scenes are acquired with a monocular camera calibrated beforehand. The features of the road lanes are extracted and recognized, and then the external parameters of the camera are calibrated by three mutually non-coincident parallel lines constructed by road lanes which have same vanishing point and different slopes on imaging plane. To compensate the error of the distance estimation caused by pitch angle, the data regression modeling method is applied, which is based on calibrating some fixed distances to image pixel samples under different pitch angles accurately. Distance of the preceding vehicle is compensated dynamically by the data regression modeling. A statistical base of 200 video highway road images in the daytime is tested in our experiments. The experimental results show that the distance estimation error ratio is lower than 5%.

Keywords: Driver Assistance System, Camera Calibration, Monocular Vision Distance Measurement, Error Compensation

1. Introduction

A main task of driver assistance vision systems is to determine the distance of the preceding vehicle. Usually it needs to know the angle and the distance of preceding vehicles in order to localize, do anti-collision warning judgment or some high-level planning[1][2]. To determine the distance to a preceding vehicle of unknown size is possible using the knowledge about the height and pitch angle of the camera and the car bearing to the point where the vehicle meets the ground. This bearing is given by the position of the vehicle in the image and the known orientation of the camera relative to the ground. So it must calibrate the camera beforehand to get the orientation. Unfortunately this orientation, especially the pitch angle, could change a lot when the vehicle runs on undulated road. As a result the calculation of the distance of a preceding vehicle to the camera usually becomes inaccurate. In this paper we show how to determine dynamically the orientation of the camera using three mutually non-coincident parallel lines and how to calculate the distance to preceding vehicles using data regression modeling.

There has been extensive work on the calibration of camera parameters. Typical intrinsic and extrinsic parameters of cameras are inferred by using specially crafted calibration objects[3][4]. Meanwhile, trilinear method has developed of the feature that without iteration and optimization calculation, expression of the extrinsic parameters of a camera, vanishing point and imaging slopes can be established through mathematical derivation according to pin-hole imaging model[5-7]. Based on the intrinsic and extrinsic parameters of cameras, coordinate transformation method, geometric derivation method and data regression modeling method are usually used to calculate vehicle distance. The first two methods are complex and their accuracy depends on the intrinsic and extrinsic parameters. So it easily leads to considerable error. As far as to the third method, its principle is simple, convenient and practical. The method accurately calibrates a few distance samples, then applies data regression modeling to build the distance detection modeling[8][9]. The method solves the affections about imaging modeling, imaging system errors, lens distortion, etc. In addition, it can adjust the fitting precision according to the actual requirements. Vision-based vehicle detection systems use various
cues that can be used for vehicle distance estimation. In contrast to these methods, our approach focuses on determining the camera pose relative to the ground during the process of moving car. While intrinsic parameters do not change after the camera is setup on the test car, the extrinsic parameters of the camera are usually hard to determine using proprioception in a highly dynamic environment like highway. But there are cues including fitting lines of lanes, lane width, and point of contact of a vehicle and the road [10-12].

2. Camera calibration

2.1. Intrinsic parameter calibration

Camera calibration and pose estimation are major issues in computer vision since they are related to many vision problems. Camera calibration consists in the estimation of a model for an un-calibrated camera. The objective is to find the internal parameters of the camera (principal point or image center, focal length and distortion coefficients), and the external parameters (position and orientation relative to a world coordinate system). One of the frequently used camera calibration techniques is the one proposed by Zhang [3]. Its implementation needs only 2D pixels in the image. According Zhang’s method, we get the camera internal parameters firstly.

2.2. Extrinsic parameter calibration

The external parameters of the camera are used to describe the relationship between camera coordinate system and world coordinate system; it shows the position and orientation of the camera in world coordinate system. The coordinate systems are established as Figure 1.

![Figure 1. The relationship between camera coordinate system and world coordinate system](image)

Practically, three mutually non-coincident parallel lines $l_1$, $l_2$, $l_3$ on the ground are used to calibrate external parameters. All the lines are parallel to the axis of the car $X_F$, and the distances between each line and the axis are $a_1$, $a_2$, $a_3$ respectively. Let’s suppose the intersection point of the three lines is $P_0(i_0, j_0)$. Selecting any point $P_k(i_k, j_k)$ on the lines except $P_0$, we have:
\[
\begin{align*}
\psi &= \arctan \left[ \frac{(r_1-r_2)(a_1-a_2)-(r_1-r_2)(a_1-a_3)}{(r_1-r_2)(r_1-a_2)-(r_1-r_2)(r_1-a_1)} \right] \\
\theta &= \arctan \left[ \frac{u_n \sin \psi / f d_n + V \cos \psi / f d_n}{h_d} \right] \\
\phi &= \arctan \left[ \frac{\cos \theta (u_n / f d_n - \sin \psi \cos \theta)}{\cos \theta} \right] \\
\varphi &= \arctan \left[ \frac{\cos \theta (u_n / f d_n - \sin \psi \cos \theta)}{\cos \theta} \right] \\
\end{align*}
\]

where

\[
\begin{align*}
A &= r_1 \sin \psi \cos \theta - \cos \psi \cos \theta \\
B &= -(\cos \phi \sin \psi + \sin \phi \cos \psi \sin \theta) - r_2 (\cos \phi \cos \psi - \sin \phi \sin \psi \sin \theta) \\
C &= r_2 \sin \psi \cos \theta - \cos \psi \cos \theta \\
D &= -(\cos \phi \sin \psi + \sin \phi \cos \psi \sin \theta) - r_2 (\cos \phi \cos \psi - \sin \phi \sin \psi \sin \theta) \\
r_n &= -(f_j / f_f)(j_n - j_a) / (j_k - j_a), \quad n = 1, 2, 3
\end{align*}
\]

\( (1) \)

2.3. Error of extrinsic parameter calibration

Equation (1) respectively for the two cases that vehicle’s position coordinates did not exceed or exceed the intersection point of cameras optical axis and ground. Due to the camera's tilt angle is generally less than 5°, the intersection point P of camera’s optical axis and ground coordinate in the world coordinate approximate to infinity. Therefore, distance of preceding vehicles calculated by equation (1) usually has relatively large errors.

It has been known that the distance of preceding vehicles is mostly depended on the tile angle of camera [13], so the precision of the calibration by non-coincident parallel lines is important for the distance measurement. The horizontal angle \( \varphi \) and direction angle \( \psi \) of the camera are adjusted to a small angle. As shown in Figure 2, camera is mounted at a height \( h \), tilt angle can be calculated by:

\[
\theta = \arctan (h / D)
\]

\( (3) \)

**Figure 2.** Camera setup for error estimation of non-coincident parallel lines calibration method

Error of non-coincident parallel lines calibration method is estimated by comparing with angles measured by laser total station. Firstly camera is installed on a tripod and the base of the tripod is adjusted to leveling. A surveyor's pole perpendicular to the ground is placed before the camera and there points A,B,C with the height of \( D_1,D_2,D_3 \) are marked out on the pole. We can get the distance \( h \) of camera optical center to the ground and the distance \( D \) of the camera to the pole. Projection position \( O \) on the ground of the camera is fixed by a plumb. Then images are captured by the camera and principal point is coincided sequentially with A, B, C on the pole by adjusting the camera. Meanwhile, camera attitude parameters \( \theta, h \) are estimated at every point. Secondly, the camera is installed on the laser total station which is adjusted to leveling. Projection position on the ground of the camera is still
the $O$ point and the distance of camera optical center to the ground is also $h$. Aim at point A from the telescope of the laser total station and recording the vertical angle $\beta_1$. In the same way, $\beta_2, \beta_3$ correspond to points B, C are recorded.

According to the triangular relationship, we have

$$\theta_i' = \arctan\left(\frac{h - D_i}{D_i}\right), \quad i=1,2,3$$

Lastly, calibration accuracy of non-coincident parallel lines calibration method is estimated by comparing $\theta_i$ and $\theta_i'$, $h_i$ and $h'_i$, $(\theta_i - \theta_i)$, $(\beta_i - \beta_i)$ and $(\theta_i - \theta_i')$ and $(\beta_i - \beta_i')$. A set of data is shown as Table 1 when $D$ is 15.10m. Taking the data of laser total station as the actual value, it can be seen that the difference accuracy of non-coincident parallel lines calibration method compared to laser total station measurement result is less than 0.1° which can satisfy the distance estimation of the preceding vehicle.

<table>
<thead>
<tr>
<th>Point A</th>
<th>$\theta_1 = 0.678^\circ$, $h_1 = 122.04\text{cm}$</th>
<th>$\theta_1' = 0.436^\circ$</th>
<th>$\beta_1 = 270^\circ 20' 78''$</th>
<th>$D_1 = 136\text{cm}$, $h = 124.5\text{cm}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point B</td>
<td>$\theta_2 = -2.322^\circ$, $h_2 = 122.42\text{cm}$</td>
<td>$\theta_2' = -2.465^\circ$</td>
<td>$\beta_2 = 267^\circ 30' 43''$</td>
<td>$D_2 = 59.5\text{cm}$, $h = 124.5\text{cm}$</td>
</tr>
<tr>
<td>Point C</td>
<td>$\theta_3 = -3.959^\circ$, $h_3 = 124.7\text{cm}$</td>
<td>$\theta_3' = -3.959^\circ$</td>
<td>$\beta_3 = 265^\circ 59' 32''$</td>
<td>$D_3 = 20.0\text{cm}$, $h = 124.5\text{cm}$</td>
</tr>
<tr>
<td>A-B</td>
<td>$(\theta_1 - \theta_2) = 3.000^\circ$</td>
<td>$(\theta_1' - \theta_2') = 2.901^\circ$</td>
<td>$(\beta_1 - \beta_2) = 2^\circ 50' 35''$</td>
<td></td>
</tr>
<tr>
<td>A-C</td>
<td>$(\theta_1 - \theta_3) = 4.445^\circ$</td>
<td>$(\theta_1' - \theta_3') = 4.395^\circ$</td>
<td>$(\beta_1 - \beta_3) = 4^\circ 21' 36''$</td>
<td></td>
</tr>
</tbody>
</table>

### 3. Estimation of the distance to the preceding vehicle

#### 3.1. Data regression modeling

The position of the footprint of the detected proceeding vehicle on the road and the parameters of the camera is utilized to estimate the longitudinal distance to the preceding vehicle. Distance of preceding vehicle is related to the image captured by the camera being installed. So we apply data regression modeling in order to estimate the distance between the preceding vehicle and the test car from the captured image. It must be pointed out that the distance includes longitudinal distance and lateral distance. Longitudinal distance is only related to pixel height of the target in the image and lateral distance is related to the pixel width of the target when the height is fixed. Therefore, the functional relationship between longitudinal distance and pixel height of the target and the relationship between lateral distance and the pixel width of the target are needed to fit. Practically, camera is installed behind the windshield of the test car and by adjusting the pitch angle $\theta$, optical axis of camera is pointed to the front 50m. Then three parallel lines are marked on the ground which parallel to the car axis and the distances of every two lines are measured. By now, camera pose can be estimated by non-coincident parallel lines calibration method. According to the calibration result, camera pose is adjusted until horizontal angle $\phi$ and direction angle $\psi$ are all less than 0.002 rad. At this time, the actual longitudinal and lateral distances of the minimum sight are measured. At every 5m between 10m to 120m along the vertical direction, markers are placed and the pixel height is recorded. At every 1.5m along the horizontal direction, markers are placed and the pixel width is recorded, which can be seen in Figure 3. Lastly, the data recorded is used to fit the relationship between the actual distance and pixel distance. Repeat the above process when optical axis of camera is pointed to the front 40m and 75m. In those three cases, pitch angle $\theta$ has been calculated by non-coincident parallel lines calibration method are $\theta_{40} = 1.712^\circ$, $\theta_{75} = 1.482^\circ$, $\theta_{75} = 1.046^\circ$. Data recorded in Figure 4 shows that there exists a tendency of slow changes at the start and rapid changes at the end.
Figure 3. Optical axis of camera is pointed to the front 50m: (a) 15m of longitudinal distance and 1.5m of lateral distance, (b) 35m of longitudinal distance and 5m of lateral distance

Figure 4. Data of image pixel distance corresponding to actual distance: (a) Relationship between horizontal distance of a pixel and vertical pixel height. (b) Relationship between actual vertical distance and vertical pixel height: (X40,Y40), (X50,Y50) are segmented to two parts (X40L,Y40L), (X50R,Y50R). (X70,Y70) are segmented into three parts: (X70L,Y70L), (X70M,Y70L), (X70R,Y70R).

The following piecewise fitting function is used to establish the relationship between actual proceeding vehicle distance and pixel distance.

\[
D_k(y) = \begin{cases} 
  a_1y^3 + a_2y^2 + a_3y + a_4 & 0 \leq y \leq y_1 \\
  b_1y^3 + b_2y^2 + b_3y + b_4 & y_1 < y \leq y_2 \\
  c_1y^{\frac{d_1}{h}} + c_2y^{\frac{d_2}{h}} & y_2 < y \leq y_3 
\end{cases}
\]

(5)

3.2. Distance compensation model

It must be pointed out that all the data acquired is under static process but bump would occur when the car is running which will change the pitch angle $\theta$ and influence the estimated distance accuracy. So we use the result of the road lane detection to compensate the distance estimated. Left and right road lane lines are as follows:

\[
\begin{align*}
  y &= k_1x + b_1 \\
  y &= k_2x + b_2
\end{align*}
\]

(6)
And the bisector line of the two lines is:

\[ y = \frac{k_1 + \lambda k_2}{1 + \lambda} x + \frac{b_1 + \lambda b_2}{1 + \lambda} \]  
\[ (7) \]

where \( \lambda = \sqrt{(1 + k_1^2)(1 + k_2^2)} \).

Three parallel lines are constructed and distances between every line to the line of the car center axis \( dl, dm, dr \) can be calculated. Vanishing point also can be calculated. Then non-coincident parallel lines calibration method is used to calculate the pitch angle \( \theta \).

Let \( VD(y), HD(y) \) are the functional relationship between pixel height and longitudinal and lateral distances, where pitch angles \( \theta_l, \theta_m, \theta_H \) are 40, 50, 75. Meanwhile \( \theta_r \) is the real-time calculated pitch angle. \( P_i \) is the horizontal resolution of the camera. \((x, y)\) is the coordinate of the preceding vehicle target. The compensated distance of preceding vehicle is:

\[ D = \sqrt{VD^2 + HD^2} \]

where:

\[ VD = \begin{cases} 
VD_{al}(y), & \theta_l \leq \theta_L \\
\frac{\theta_u - \theta}{\theta_u - \theta_l} VD_{al}(y) + \frac{\theta_l - \theta}{\theta_l - \theta_u} VD_{al}(y), & \theta_L < \theta < \theta_m \\
\frac{\theta_u - \theta}{\theta_u - \theta_m} VD_{al}(y) + \frac{\theta_m - \theta}{\theta_m - \theta_u} VD_{al}(y), & \theta_m < \theta \leq \theta_H \\
VD_{al}(y), & \theta > \theta_H 
\end{cases} \]

\[ (9) \]

\[ HD = \begin{cases} 
\text{abs}(x - \frac{P_c}{2}) \cdot HD_{al}(y), & \theta_l \leq \theta_L \\
\text{abs}(x - \frac{P_c}{2}) \cdot \frac{\theta_u - \theta}{\theta_u - \theta_l} HD_{al}(y) + \frac{\theta_l - \theta}{\theta_l - \theta_u} HD_{al}(y), & \theta_L < \theta \leq \theta_m \\
\text{abs}(x - \frac{P_c}{2}) \cdot \frac{\theta_m - \theta}{\theta_m - \theta_l} HD_{al}(y) + \frac{\theta_l - \theta}{\theta_l - \theta_m} HD_{al}(y), & \theta_m < \theta \leq \theta_H \\
\text{abs}(x - \frac{P_c}{2}) \cdot HD_{al}(y), & \theta > \theta_H 
\end{cases} \]

\[ (10) \]

4. Experimental results

The proposed system was implemented on the TI TMS320DM6437 (600MHz system clock) embedded hardware platform with a few peripheral devices (e.g., a LCD screen, Ethernet, AC-97, composite video interface, etc.). A CMOS camera with PENTAX 8.5mm 1:1.5 TV lens was connected to the hardware platform with a composite video interface and mounted behind the windshield of the test car to acquire a QVGA (320 by 240) resolution image at 30 frames. Camera internal parameters were calibrated by Zhang’s method firstly. We can get that the equivalent focus of the lens \((f_i, f_j)\) is \((1686.09, 1690.13)\), the principal point is \((319.5, 239.5)\). Additionally, the intrinsic matrix is \([1686.0925, 0.0000, 319.5000; 0.0000, 1690.1278, 239.5000; 0.0000, 0.0000, 1.0000]\). And distortion vector is \([-0.3147, 12.5536, -0.0042, -0.0056]\), rotation vector is \([-2.0512, -2.1521, 0.3275]\), translation vector is \([-27.3676, -32.3330, 338.9515]\). According to the data captured, we get the fitted coefficient of equation (2), the result is shown in Figure 5.

The experiments were performed in the daytime with sufficient natural light. Preceding vehicle stopped at a serial fixed distances, the distances being compensated are shown in Figure 6 and Table 2. Distances between test car center axis and right lane are shown in Figure 7 and Tab.3. A statistical base.
of 200 vehicle video images is tested in our experiments; the experimental results show that the relative error of longitudinal distance is less than 2.9% and relative error of lateral distance is less than 2.07%. On the basis of the experimental results, we say that the proposed distance estimated method can satisfy the demand of driver assistance system in real time.

![Graphs showing the relationship between image pixel distance and actual distance.](image)

**Figure 5.** Data of image pixel distance corresponding to actual distance: (a) Relationship between vertical distance of a pixel and vertical pixel height. (b) Relationship between actual horizontal distance and vertical pixel height.

![Images showing vehicle at different distances.](image)

(a) 20m  (b) 40m  (c) 60m
Figure 6. Preceding vehicle distances estimated from 20 to 120m

<table>
<thead>
<tr>
<th>Actual distances(m)</th>
<th>Estimated distances(m)</th>
<th>Difference(m)</th>
<th>Relative error</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>19.1</td>
<td>-0.9</td>
<td>-4.71%</td>
</tr>
<tr>
<td>30</td>
<td>28.9</td>
<td>-1.1</td>
<td>-3.81%</td>
</tr>
<tr>
<td>40</td>
<td>39.3</td>
<td>-0.7</td>
<td>-1.78%</td>
</tr>
<tr>
<td>50</td>
<td>48.5</td>
<td>-1.5</td>
<td>-3.09%</td>
</tr>
<tr>
<td>60</td>
<td>59.2</td>
<td>-0.8</td>
<td>-1.35%</td>
</tr>
<tr>
<td>70</td>
<td>66.9</td>
<td>-3.1</td>
<td>-4.63%</td>
</tr>
<tr>
<td>80</td>
<td>77.3</td>
<td>-2.7</td>
<td>-3.49%</td>
</tr>
<tr>
<td>90</td>
<td>91.8</td>
<td>1.8</td>
<td>1.96%</td>
</tr>
<tr>
<td>100</td>
<td>101.2</td>
<td>1.2</td>
<td>1.19%</td>
</tr>
<tr>
<td>110</td>
<td>112.3</td>
<td>2.3</td>
<td>2.05%</td>
</tr>
<tr>
<td>120</td>
<td>125.4</td>
<td>5.4</td>
<td>4.31%</td>
</tr>
</tbody>
</table>

Figure 7. Distances between test car center axis and right lane
Table 3. Lateral distance error

<table>
<thead>
<tr>
<th>Actual distances(cm)</th>
<th>Estimated distances(cm)</th>
<th>Difference(cm)</th>
<th>Relative error</th>
</tr>
</thead>
<tbody>
<tr>
<td>58</td>
<td>52</td>
<td>-6</td>
<td>-10.34%</td>
</tr>
<tr>
<td>71</td>
<td>70</td>
<td>-1</td>
<td>-1.41%</td>
</tr>
<tr>
<td>90</td>
<td>93</td>
<td>3</td>
<td>3.33%</td>
</tr>
<tr>
<td>102</td>
<td>103</td>
<td>1</td>
<td>0.98%</td>
</tr>
<tr>
<td>113</td>
<td>111</td>
<td>-2</td>
<td>-1.77%</td>
</tr>
<tr>
<td>129</td>
<td>128</td>
<td>-1</td>
<td>-0.78%</td>
</tr>
<tr>
<td>142</td>
<td>144</td>
<td>2</td>
<td>1.41%</td>
</tr>
<tr>
<td>157</td>
<td>158</td>
<td>1</td>
<td>0.64%</td>
</tr>
<tr>
<td>166</td>
<td>165</td>
<td>-1</td>
<td>-0.60%</td>
</tr>
<tr>
<td>181</td>
<td>182</td>
<td>1</td>
<td>0.55%</td>
</tr>
<tr>
<td>215</td>
<td>213</td>
<td>-2</td>
<td>-0.93%</td>
</tr>
</tbody>
</table>

5. Conclusion

This work presented a monocular vision distance measurement method based on dynamic error compensation in our highway driver assistance system that can be implemented in a cost-effective embedded system with limited hardware resources. The proposed method is successfully implemented on a test car and tested on Highway in Tianjin, and its effectiveness is verified. The results of simulation and real vehicle test show that the compensation algorithm can effectively reduce distance measurement errors of the preceding vehicle caused by pitch angle changes when the vehicle runs on undulated road.

6. Acknowledgement

The present work was supported by the Tianjin Science and Technology Key Project under grant No. 10ZCKFGX00300, and in part by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry.

References


