Research on vehicle camera calibration and ranging technology based on texture features

Cai Wang1, Bifeng Cui1*, Yanxu Zhu1, Yuhang Zhang1, Qixuan Li1 and Zhao Wei1

1 Information, Beijing University of Technology, Beijing, Beijing, 100124, China

*Corresponding author’s e-mail: cbf@bjut.edu.cn

Abstract. In order to realize the ranging function of the vehicle monocular camera on unstructured roads, it is proposed to use the method based on texture features and soft voting to detect the vanishing point of the road and the road edge, and use camera calibration method and tri-linear method to obtain the internal and external parameters of the camera. Obtain the distance information in the image according to the geometric relationship model. Experiments show that this method can measure the distance of an unstructured road image with only the centerline of the road relying only on the vehicle monocular camera. The method has a wider application range and is simpler in actual engineering.

1. Introduction

Vehicle speed is strictly restricted in road traffic, standard structured roads are generally equipped with dedicated speed-measuring devices, but the number of such devices is relatively small, and they often play little role in the determination of the liability for traffic accidents. At present, vehicle monocular camera video computing has become one of the common methods for estimating the speed of vehicles in traffic accidents in China’s judicial appraisal.[1]

Traditional calculations method of Vehicle speed in traffic accident video is mainly based on the method of body reference and ground reference[2], however, this method has a large number of human factors and a high error range. In order to make the speed measurement results more objective, in the aspect of monocular vision ranging and speed measurement, Chan[3] et al. proposed a method of estimating the distance using the height of the character of the front license plate, and obtaining the height of the character pixel through license plate positioning, tracking and optical recognition. Park[4] et al. obtained the distance information of the front vehicle through the geometric model according to the change of the pitch angle of the visual sensor caused by the road tilt. Guo Lei[5] established the mapping relationship between the image coordinate system and the world coordinate system on the basis of the perspective projection principle, and then obtained the longitudinal distance information between the visual sensor and the front obstacle. Liu Jun[6] et al. established a lateral and longitudinal distance ranging model of the front vehicle based on the geometric model. This experiment shows that the lateral and longitudinal ranging errors do not exceed 5%.

In general, the current research results of vehicle monocular camera ranging are mainly used for vehicle collision warning, and the detection targets are often vehicles, lane lines in structured roads, etc. For unstructured roads, it is difficult to achieve distance measurement without relying on the support of sensors. In this paper, a method based on estimating the texture direction is used to obtain the area and vanishing point of unstructured roads. The tri-linear method is used to calibrate the external parameters of the camera, and combined with the distance measurement model of the
geometric relationship, the vertical distance from a point in front of the camera to the vehicle is obtained. The method can realize the distance measurement of the vehicle monocular camera on the unstructured road. The method has a wider application range and is easier to apply in actual engineering.

2. Ranging model
The ranging model mainly includes vanishing point and road area detection, camera parameter calibration and ranging model.

2.1. Vanishing point and road area detection
Extracting the depth information of monocular visual images depends on the detection of vanishing points. For video files obtained by cameras at different heights and different pitch angles, the vanishing point needs to be extracted according to the specific situation of its image. Therefore, the automatic detection of vanishing point is the focus of ranging technology. In this paper, the Gabor filter is first used to calculate the texture direction of each pixel in the road image. For direction and scale, the gabor filtering kernel function is expressed by formula (1):

\[
\Psi_{\omega,\phi}(x,y) = \frac{\omega}{\sqrt{2\pi\epsilon}} e^{-\frac{1}{8\epsilon^2} \left(4x^2 + y^2 \right)} \left( e^{i\omega \phi} - e^{-\frac{\omega^2}{2}} \right)
\]

\(I(x,y)\) is a point on a grayscale image, the convolution of image \(I\) and the Gabor kernel under direction \(\phi\) and scale \(\omega\) is defined as

\[
G_{\omega,\phi} = I \otimes \Psi_{\omega,\phi}
\]

In direction \(\phi\), the response image \(R_{\phi}(z)\) is calculated by the average value of the square modulus of \(G_{\omega,\phi}\) at different scales, As shown in formula (3):

\[
R_{\phi}(z) = \text{average}_\omega \{ (\text{Re}(G_{\omega,\phi}))^2 + (\text{Im}(G_{\omega,\phi}))^2 \}
\]

According to the average maximum response image, the texture direction angle \(\theta(z)\) is defined as follows[8]:

\[
\theta(z) = \arg \max_{\phi} R_{\phi}(z)
\]

The effect picture of the pixel texture direction of the image is shown in Figure 1 (select only some pixels), the scale is selected as 5, and the direction is selected from 0 to 180 degrees, with 5 degrees as the interval.

![Figure 1 texture direction rendering of the image](image-url)
After completing the calculation of the texture direction, the vanishing point can be detected according to the local soft voting method.

The voting scores of all pixels in the candidate area are calculated by the following soft voting function[4]:

$$\text{Vote}(P,V) = \frac{\exp(-\lambda / \gamma)}{1 + d(P,V)^\gamma}$$  \hspace{1cm} (5)

Where \( V \) is a pixel (candidate point) in the vanishing point candidate area of point \( P \), \( \lambda \) is the distance from point \( V \) to the pixel with the same \( y \) coordinate on \( PQ \). \( \gamma = 180 \) is a fixed parameter, \( d(P,V) \) is the quotient of the distance between \( P \) and \( V \) and the diagonal length of the image.

The voter candidate area is shown in Figure 2, (Area \( p_1, i_1, j_1 \) in Figure 2). \( P \) is the voter (the pixel of the vote), \( \overrightarrow{OP} \) is the direction, \( \epsilon \) is the angle area (\( \epsilon = 5^\circ \)), \( |PK| = 0.50 \times H \), \( |PQ| = 0.65 \times H \), where \( H \) is the height of the picture.

![Figure 2 VP candidate area](image)

The voting and scoring results are shown in Figure 3:

![Figure 3 effect of voting results](image)

After determining the coordinates of the vanishing point in the figure, the road area can be detected using the road area detection method based on the vanishing point[8]. In this method, the histogram can be calculated by the Angle difference (AD) and color difference (CD) of each voting pixel, and the boundary between two roads can be detected by using the histogram. Calculate the value of the histogram according to the following formula:

$$h(u) = \sum_{i=1}^{N} \text{diff}(P_i) \times \text{diff}(R_{i1}, R_{i2}) \times \delta[u - \alpha_i]$$  \hspace{1cm} (6)

Where \( N \) is the number of voters, \( u = 0, 1, ..., 180 \) , \( \delta \) is a Kronecker symbolic function.
Exclude angles less than 20 degrees and greater than 160 degrees in the histogram.

After the histogram is generated, the road boundary can be detected by selecting the two angles with the largest value in the histogram. An example of using the histogram to detect the two road boundaries is shown in Figure 4.

![Figure 4 Angle histogram and road area detection diagram](image)

### 2.2. Calibration of camera internal and external parameters

For the calibration of internal parameters, Zhang's camera calibration method\[11\] has been very mature in actual engineering applications, the corresponding program set and toolbox are relatively complete, and the requirements for the user’s knowledge base of computer vision are low, therefore, the algorithm and derivation process not be repeated here.

The tri-linear method based on Zhang's camera calibration method is used for camera external parameter calibration. The tri-linear method is derived based on the fact that three random parallel lines in the image have the same vanishing point and different slopes. Li Qing\[12\] et al. gave a detailed derivation of the method. The tri-linear method only requires three parallel lines on the ground, then according to the intersection and slope of the three straight lines, the external parameters of the camera can be calibrated.

The specific algorithm is as follows:

\[
\psi = \arctan \left( \frac{(r_1 - r_2)(a_i - a_2) - (r_1 - r_2)(a_1 - a_i)}{(r_1 - r_2)(r_1a_i - r_2a_2) - (r_1 - r_2)(r_2a_1 - r_1a_2)} \right),
\theta = \arctan \left( \frac{u_i \sin \psi}{f/d_x} + \frac{v_a \cos \psi}{f/d_y} \right)
\]

\[
\phi = \arctan \left( \frac{\cos \theta (u_a / f_d) + \sin \psi \tan \theta}{\cos \psi} \right)
\]

\[
h = \frac{(a_2 - a_1)AC}{BC - AD}
\]

\[
d = \frac{B}{A} \frac{AC}{BC - AD} + a_i
\]

Where

\[
A = r_1 \sin \psi \cos \theta - \cos \theta \cos \psi
\]

\[
C = r_1 \sin \psi \cos \theta - \cos \theta \cos \psi
\]

\[
B = - (\sin \psi \cos \theta + \sin \psi \cos \psi \sin \theta) - r_1 (\cos \psi \cos \psi - \sin \psi \sin \psi \sin \phi)
\]

\[
D = - (\sin \psi \cos \theta + \sin \psi \cos \psi \sin \theta) - r_2 (\cos \psi \cos \psi - \sin \psi \sin \psi \sin \phi)
\]

\[
r_n = -(f_j / f_i) (i_h - i_n) / (j_h - j_n), n = 1, 2, 3
\]

\[i_n, j_n\] is the coordinate of the pixel coordinate system on the nth straight line, \[i_h, j_h\] is the pixel
coordinate system coordinates of the vanishing point, \( f_i \), \( f_j \) are the respective focal lengths in the \( i \) direction and \( j \) direction, obtained from the internal parameters in 2.2.

### 2.3. Ranging algorithm

The ranging algorithm based on the projection model of the geometric relationship\(^{(7)}\) describes the principle of perspective transformation based on the small hole imaging model. According to geometric relations, obtain the horizontal distance \( d \) between a point \( P \) in the image and the center of the lens.

As shown in Figure 5, \( f \) is the effective focal length of the CCD camera, \( \alpha \) is the pitch angle of the CCD camera, \( h \) is the installation height of the CCD camera (the height from the center of the lens to the ground), \((x_0, y_0)\) is the intersection point of the optical axis and the image plane, as the origin of the image physical coordinate system, \((x, y)\) is the projected coordinate of a point \( P \) on the road on the image plane. According to the geometric relationship, the calculation formula of the horizontal distance \( d \) between the point \( P \) and the lens center can be obtained:

\[
d = h / \tan \{\alpha + \arctan[(y_0 - y) / f]\}
\]  

\[\text{(12)}\]

![Figure 5 (a) Camera layout drawing (b) Distance geometry model](image)

### 3. Calibration of camera external parameters

#### 3.1. Camera internal parameter calibration experiment

The internal parameters of the camera are completed by Zhang's calibration method \(^{(11)}\), and the calibration platform is built by EmguCV3.0 and .Net development environment. The calibration results of CCD camera information and internal parameter are shown in Table 1.

| Parameter                  | Value         | Unit       |
|---------------------------|---------------|------------|
| Camera model:             | Sony IMX519   |            |
| Sensor size:              | 1/2.6         | inch       |
| Unit pixel area           | 1.22          | micron     |
| \( c_i \)                 | 2132.53       |            |
| \( c_j \)                 | 1461.78       |            |
| \( f_i \)                 | 3308.68       |            |
| \( f_j \)                 | 3320.42461    |            |
| \( d_x \)                 | 0.00105969125 | mm/pix     |
| \( d_y \)                 | 0.00105969125 | mm/pix     |
3.2. Calibration of camera external parameters
The camera external parameter calibration is completed by the aforementioned tri-linear method, the calibration platform is built by the .Net development environment, take the optical center of the camera as the origin of the vehicle body coordinate system. The vanishing point and the pixel information on both edges of the road required by the three-line method are obtained from 2-1, the road centerline information is extracted from the road centerline. For the calibration diagram of unstructured road shown in Figure 6, since there is no road centerline, a straight line can be fitted from the middle seam caused by road construction to obtain the pixel information of the third line.

![Figure 6 Diagram of three-line method on-site calibration](image)

The camera external parameters are calculated by the formula (7)–(10) in 2.2 according to the distance between the vehicle body coordinate system and the three straight lines, the vanishing point and the pixel coordinates of the three points on the three straight lines. The calibration results are shown in Table 2.

| Calibration value | Psi 0.033514 rad | Theta 0.05291 rad | Phi -0.02435 rad | d -100.5254 mm | h 1202.6617mm |
|-------------------|------------------|-------------------|------------------|----------------|----------------|

3.3. Ranging experiment
As shown in Figure 7, the pixel coordinates of objects at different distances are extracted, and the distance between the vehicle and the object in front is calculated according to the formula (12) in 2.3.

![Figure 7 Pixel coordinates of objects at different distances](image)

(a) The distance is 7403mm          (b) The distance is 11102mm       (c) The distance is 14809mm

The ranging results and errors are shown in Table 3.
Table 3 Ranging experiment results

| Actual distance | Experimental ranging value | Accuracy  |
|-----------------|----------------------------|-----------|
| 3704mm          | 3488.810mm                 | 94.2922% |
| 7403mm          | 6908.129mm                 | 93.3351% |
| 11102mm         | 9940.786mm                 | 89.5566% |
| 14809mm         | 12935.901mm                | 87.4047% |
| 18512mm         | 16621.504mm                | 89.846%  |
| 22196mm         | 20116.345mm                | 90.6142% |

Carry out 10 sets of ranging experiments in different road conditions according to the above steps. Experiments show that for ordinary rural dirt roads under general conditions, the error of the vanishing point and road area detection is large, which makes it difficult to complete the external parameter calibration experiment. However, on ordinary unstructured cement roads with middle seams or urban structured roads with only a single lane line, the calibration error of this method can generally be controlled within 20%. The corresponding analysis results are shown in Table 4.

Table 4 Ranging effect of this method under different road conditions

| Road area detection effect | racetrack | Cement road 1 | Cement road 2 |
|---------------------------|-----------|---------------|---------------|
| Error                     | < 12%     | < 15%         | < 20%         |

Compared with the methods in literature [3]-[6], the method proposed in this article can be applied to ordinary cement roads with middle seams, urban roads with only centerline, and standard urban structured roads. This method does not need to detect vehicle, lane lines and other information, nor does it need to estimate changes in pitch angle of the camera.

4. Conclusion
This paper applies the road vanishing point and road area detection methods based on texture features to camera calibration and road ranging model, which can reduce the dependence on lane lines, zebra crossings, vehicles and other target information in the ranging process. It helps increase the robustness of the ranging model and the application range in actual engineering, and provides a further theoretical reference for the automation of the calibration process. Experiments show that this method has a good effect on the ranging of unstructured roads. It can be performed on ordinary cement roads with middle seams and roads with only the centerline. The experimental platform is relatively simple to build.

In terms of system errors, first of all, due to the large amount of calculation in the texture voting process, the performance requirements of the computing platform are relatively high. Therefore, in the experiment process, the picture had to be compressed in advance, and then the coordinate conversion was performed, which caused a certain detection error. Therefore, the calculation speed of texture voting still needs to be further optimized. Secondly, in the geometric ranging model, the pitch angle of camera has a greater impact on the ranging results, so in the tri-linear method calibration process, multiple calibrations can be performed by selecting multiple sets of points on three lines, the final result is obtained after data processing.
References

[1] GB/T33195-2016, Vehicle Speed Appraisal in Road Traffic Accidents[S]. Beijing: China Standard Press, 2016.

[2] Guan Chuang. Research on speed identification of forward vehicles based on image perception technology[D]. Chang'an University, 2019.

[3] KinYip Chan, Andrzej Ordys, Olga Duran. A System to Measure Gap Distance between Two Vehicles Using License Plate Character Height[C]// International Conference on Computer Vision & Graphics. Springer Berlin Heidelberg, 2010.

[4] Ki-Yeong P, Sun-Young H. Robust Range Estimation with a Monocular Camera for Vision-Based Forward Collision Warning System[J]. Scientific World Journal, 2014, 2014:1-9.

[5] Guo Lei, Xu Youchun, Li Keqiang, et al. Research on real-time ranging method based on monocular vision[J]. Journal of Image and Graphics, 2006(01):77-84.

[6] Liu Jun, Gao Xueling, Wang Liming, et al. Research on Forward Vehicle Detection and Forward Collision Warning Algorithm Based on OpenCV[J]. Automobile Technology, 2017, 000(006):11-16.

[7] Wang Y, Teoh E K, Shen D. Lane detection and tracking using B-Snake[J]. Image and Vision Computing, 2004, 22(4):269-280.

[8] Trung Hieu BUI,Eitaku NOBUYAMA,Takeshi SAITO. A Texture-Based Local Soft Voting Method for Vanishing Point Detection from a Single Road Image[J]. The Institute of Electronics, Information and Communication Engineers, 2013, E96.D(3).

[9] Kong Hui,Audibert Jean-Yves,Ponce Jean. General road detection from a single image[J]. IEEE transactions on image processing : a publication of the IEEE Signal Processing Society, 2010, 19(8).

[10] Bui T H, Saitoh T, Nobuyama E. Road area detection based on texture orientations estimation and vanishing point detection[C]// The SICE Annual Conference 2013. IEEE, 2014.

[11] Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11):P.1330-1334.

[12] Li Qing. Calibrate the external parameters of the vehicle camera with the three-line method[J]. Photoelectric Engineering, 2004(08):23-26.

[13] Wang Zhanqing, Chen Shunyun. Overview of Vehicle Distance Measurement Methods Based on Monocular Vision[J]. Science & Technology Information, 2010(27):39-42+44.