Zebra-Crossing Automatic Recognition and Early Warning for Intelligent Driving

Mingli Yang¹,³, Hua Shi²*, Yihui Chen⁴, Xiangjin Ran¹ and Chunhui Xu¹

¹College of Earth Sciences, Jilin University, Changchun, 130061, P.R.China
²School of Management, Shanghai University, Shanghai, 200444, P.R.China
³Department of Information Engineering, Jilin Business and Technology College, Changchun, 130507, P.R.China
⁴Department of Information Technology, ChangChun Vocational Institute of Technology, Changchun, 130033, P.R.China

shihu_shdx@sina.com

Abstract

Zebra-crossing Recognition is one of the essential parts of the visual based intelligent vehicle navigation or intelligent driving assistant system. In order to address the real-time and robustness, Zebra-crossing Recognition Method which is based on spatial-temporal correlation has been proposed. Firstly, calibrate a camera mounted on the vehicle by a practical method. Then, according to the prior knowledge such from GPS etc, a judgment whether it's in the Crossing area is made. Next, utilize the bipolar property of Zebra-crossing to extract features. Finally, the recognition results are obtained according to the model constraints. In this paper, proposed methods can improve real time identification of the zebra line by using spatial correlation, reduce the cost of recognition and lower errors during identification. The method overcomes some disadvantages of traditional identification approaches based upon video recognition, for instance higher cost and errors.

Keywords: Zebra-crossing Recognition, Temporal Correlation, Intelligent Driving

1. Introduction

A safe road-crossing system is an extreme necessity for intelligent driving to know whether there is a crossing ahead of it. The zebra-crossing often characterizes a crossing, which provides a safe area for pedestrians to cross the road reminds vehicles to reduce their speed and pays attention to safety.

Currently, zebra-crossing recognition based images has attracted researchers' attention. Crisman [1] decided whether there is Zebra crossing through the analysis of the interval between two consecutive two strips, the number and direction of it. Stephen [4] proposed the vanishing point method to identify the zebra crossing in the road. Cao [3] applied the neural network algorithm to the traffic sign recognition research which also obtained good effects.

On the basis of previous studies, I put forward a new intelligent driving Zebra-crossing Recognition Method. This method relies on temporal and spatial correlations, improves the IPM(Inverse Perspective Mapping) and uses classic Bipolar coefficient.

We concentrate on the detection of the location of a zebra-crossing by a single camera image, because necessary information is obtained using only a camera coordinate system rather than a world coordinate system.

*Corresponding Author

ISSN: 2005-4254 IJSIP
Copyright © 2016 SERSC
The recognition algorithm is outlined in Section 3. In Section 2, three methods are described to detect and estimate the pose with results. Section 3 explains the principle of the proposed technique. Method for the recognition of the existence of crossing is presented in Section 4. Section 5 presents the experimental results, followed by the conclusions and future research directions in Section 6.

2. Related Work

In the 1990s, Stephen Se [6-10] proposed a method of using the Vanishing Point. He considered the zebra line is parallel in three dimensions, and the two-dimensional zebra line which is the projection of 3D onto a camera will gather together with Vanishing Point in the end. Figure 1, shows a cuboid’s vanishing points from perspective projection. According to this theory, firstly, get approximate parallel lines with application of Hough transform, after that, find the focus again, namely Vanishing Point. Finally, identify the Zebra-cross area based on the labeled location of black and white cross.

![Figure 1. Vanishing Points from Perspective Projection Cuboid's](image)

Vanishing Point has the advantage of having a simple operation. We can obtain the intersection of straight line segments through Hough transform, and then determine whether it is a Zebra line segment via the number of crossed line segments, it can meet the requirement of real-time under the simple background. However, the acquisition of Vanishing Point is quite difficult when the pavement is deformed and the camera is distorted. In addition, the accuracy of Vanishing Point method is not high enough in the case of having an insufficient Zebra line. Having said that, Vanishing Point could be used as a reference with other methods for its simple calculation.

Bipolar coefficient method was first put forward by Mohammad in 2005, he uses the mean and variance of image to deduce a method of judging distributed intensity of black and white in the images [9-11]. It is often necessary to image block and calculate its bipolar coefficient. The ideal zebra crossing image block has perfect bipolarity, simple calculation of the bipolar coefficient as well as fast computing speed.

Since the Zebra crossing image has the characteristic of separated mutually black-and-white stripes and these stripe intervals are the same, therefore, the Frequency Domain image must have a peak on a particular frequency. That’s what Sebastian group did. First of all, they proceeded to histogram equalization on image, which largely improve the
contrast of the processed image. Then, they adopted Canny edge detector and Morphological Dilation. Finally, DFT transform was performed on each line of image.

Vanishing Point has the characteristics of fast in running but with a problem of lower discrimination. The error will be accumulated for the Zebra crossing. For images with fewer lines image and more parallel line, it is difficult to distinguish. Frequency domain transformation requires that the Zebra crossing is parallel to the photographer and the camera cannot have too much distortion. Even though this method has a high recognition rate, it still cannot meet the needs of real-time computing due to the complicated calculation. According to the characteristics of well-defined colors, bipolar coefficient method uses the mean and variance to calculate the degree of black and white in a specific area, the black and white degree is measured by the ratio of the mean and total variance. This approach possesses high computational speed and it is a good method for distinguishing Zebra crossing area.

3. Principle of the Proposed Technique

Two images of a real road scene containing pedestrian crossing is shown in Figure 2. It is clear that Zebra-crossing information is mainly focused on the lower image area in structured road [13-16].

![Figure 2. Real Road Image from Monocular Camera in Car](image)

The crossing pattern can be treated as a bipolar region. The proposed technique includes four major steps.

3.1. Monocular Camera Calibration

Internal and external parameters of the camera are two sets of parameters. They are the inside and outside parameters of the camera respectively. In the case of no distortion of the lens, the camera could be viewed as an ideal pinhole imaging. The pinhole camera model is shown in Figure 3, 0 is the projection center and 0 is the intersection point of the optical axis and the imaging plane. In the ideal case, 0 is also the center of the imaging plane.

![Figure 3. Camera Pinhole Imaging Model](image)

According to the principle of similar triangle, results are shown in the follow:

\[ x_x = f \frac{X_x}{Z_x}, \quad y_y = f \frac{Y_y}{Z_y}, \]

\[ x_{x^*}, y_{y^*} \] is the two-dimension coordinate in the \( O_{x^*} y_{y^*} \) imaging plane coordinate system;
$x, y, z, f$ is the three-dimension coordinate in the $O_x, O_y, O_z$ world coordinate system; $f$ is focal length.

The definition of Inverse perspective mapping is the process of turning the perspective images into the inverse perspective image via transformation matrix. Two coordinate systems are defined in Euclid space: $W$(world coordinate system) and $I$(image coordinate system) respectively.

$W = \{(x, y, z)\} \in E^3$  
$I = \{(u, v)\} \in E^2$

Inverse perspective transformation is shown as Figure 4.

![Figure 4. IPM Transform](image)

Put the on-board camera in the vehicle's central mirror, set the camera's plane parallel to the horizontal plane, the projection of camera's plane is the origin of the WCS(world coordinate system). The position of the on-board camera installed in the car body is $(0, 0, h)$ of the WCS, and other parameters of camera calibration are as follows:

Put the on-board camera in the vehicle's central mirror, set the camera's plane parallel to the horizontal plane, the projection of camera's plane is the origin of the WCS(world coordinate system). The position of the on-board camera installed in the car body is $(0, 0, h)$ of the WCS, and other parameters of camera calibration are as follows:

![Figure 5. (a) The $xoz$ Plane of the WCS (b) The $yoz$ Plane of the WCS](image)

The inverse perspective transformation of image coordinate system to the world coordinate system is as follows:

$$x = h \times \cot \left( \frac{2\alpha}{R_y - 1} u - \alpha \right) \times \sin \left( \frac{2\alpha}{R_x - 1} v - \alpha \right)$$
\[ y = h \times \text{ctg} \left( \frac{2\alpha}{R_y - 1} \right) \times \cos \left( \frac{2\alpha}{R_y - 1} \right) \]

\[ z = 0 \]

According to above formula, this method requires four parameters, including two parameters, camera height \( h \) and perspective \( 2\alpha \), which need to be measured manually.

### 3.2. Bipolarity

An image is obtained after preprocessing, and the gray level of the image is \( p_0(x) \). If the block contains only black and white pixels, then \( p_0(x) \) can be written as

\[ p_0(x) = \alpha \ p_1(x) + (1 - \alpha) \ p_2(x) \quad (1) \]

Where \( 0 \leq \alpha \leq 1 \), \( p_1(x) \) is the intensity distribution of black pixels and \( p_2(x) \) is the intensity distribution of white pixels.

The mean and the variance of the image gray level can be expressed as:

\[ \mu_i = \int p_r(x) \ dx \quad (2) \]

\[ \delta_i^2 = \int (x - \mu)^2 \ p_r(x) \ dx \quad (i = 0,1,2) \quad (3) \]

Using the above relations, we can write \( \delta_0^2 \) as

\[ \delta_0^2 = a\delta_1^2 + (1 - a)\delta_2^2 + a(1 - a)(\mu_1 - \mu_2)^2 \quad (4) \]

Equation (4) shows that the total variance consists of the weighted sum of variances and the difference of means. If \( \delta_0^2 = a(1 - a)(\mu_1 - \mu_2)^2 \), then \( p_0(x) \) can be said almost bipolar.

So, we define the bipolarity \( \gamma \) as

\[ \gamma = \frac{1}{\delta_0^2} (a(1 - a)(\mu_1 - \mu_2)^2) \quad (5) \]

Equation (5) implies that \( 0 \leq \gamma \leq 1 \). If \( \gamma = 1 \), there are \( \delta_1 = \delta_2 = 0 \), and

\[ p_1(x) = \delta(x - \mu_1), \ p_2(x) = \delta(x - \mu_2) \quad (6) \]

So, \( \gamma = 1 \) corresponds to perfect bipolarity and \( \gamma = 0 \) represents the absence of bipolarity.

The bipolar coefficient method usually needs to divide the image into blocks and to calculate the bipolar coefficient in each block. The ideal zebra image blocks have prefect bipolarity, and it is simple and fast for bipolar coefficient calculation. Usually, there are two methods for block partition: the uniform size and non-uniform size block-divided method. And both methods are based on the same rule: the length and width of the block cannot be less than the width of the zebra line.

### 3.3. Extraction of Zebra-Crossing Features Points

Zebra-crossing is regularly arranged, as shown in Figure 6(a). Those solid circles denote edge between white and black bands. We may use simple edge, such as Canny,
Sobel, Prewitt and so on, in detection technique to find feature points. Fisher criterion [14], however, technique gives better result as it works on local maximum feature compared to the well-known edge detection technique that works on the basis of global threshold.

Consider a horizontal window of size \((2 \times W + 1,2)\) pixels whose center coincides with a feature point as shown in Figure 6(b). We put 6 pixels into the value of win. Let Class 1 consist of pixel intensities in the left half of the window and Class 2 consist of intensities in the right half of the window. Thus, \(f_c\) is defined as the ratio of \(\delta_b\) to \(\delta_w\). \(\delta_b\) is between class variance, and \(\delta_w\) is within class variance. Then, we can write \(f_c\) as:

\[
f_c = \frac{\delta_b}{\delta_w}
\]

\[
(7)
\]

Where \(\delta_b = p_1 p_2 (m_1 - m_2)^2\), \(\delta_w = p_1 v_1 + p_2 v_2\).

\(m_i\) is mean intensity, \(v_i\) is the variance of intensity and \(p_i\) is the probability of class \(i\) \((i = 1,2)\).

Figure 6. (a) Feature Points in a Zebra-Crossing, (b) The Horizontal Window At a Feature Point

3.4. Model Constraint and Temporal Correlation

Depending on the types of junctions, there are two kinds of Zebra-crossing: orthogonal and oblique zebra line. Crossroads and T-junction are generally orthogonal zebra-lines, while other special intersection such as three-ways crossroads are usually oblique zebra-lines, as shown in Figure 7.

Figure 7. The Types of Zebra Lines

Generally, Zebra crossing is perpendicular to the center line of the Zebra crossing. Under some special circumstances, the angle of center line of the Zebra crossing and Zebra-line neither better neither less than 60 degrees nor larger than 120 degrees. Zebra lines should be parallel to the center line of the road. The width of the zebra crossing should be 300cm, and it can be increased up to 500cm at most. The height of the Zebra-line should between 40cm and 45cm, while a better interval is 50cm. The interval can be changed with the width of a road. The interval, however, should be less than 80cm. For most urban roads, the width of the white zebra line is 45cm and the interval is 60cm.
In this paper, W denotes the width of the white stripes in the zebra crossing, H is the height of the white stripes in the zebra crossing, while w is the width of the interval between stripes of zebra crossing, and N is the number of the white stripes in the zebra crossing. And they satisfy the following conditions: 40cm<W<50cm, 300cm<H<500cm, 55cm<w<65cm, 3<N<10.

On the other hand, there may be traffic lights, stop lines and other prior knowledge at a junction, making full use of this information can be used to test the correctness of the results. Because of the good quality of GPS signal, intelligent driving can acquire information of the traffic intersection through the GPS, including the type of the intersection and distance from the intersection. When the distance was less than 100 meters, zebra-crossing identification will be started promptly.

4. Zebra-Crossing Detection

A flow diagram of the proposed method is given in Figure 8.

![Flow Diagram of the Proposed Method](image)

This method is based on the high-contrast of the zebra crossing gray level information and use the bipolar property to distinguish the zebra crossing under the constraints of GPS. The process of using bipolar is shown in Figure 9:
Figure 9. Flow Chart of Zebra Line Identification Method Based on Bipolar

This paper using bipolar for image segmentation, high bipolar region is extracted as the candidate region, and fusing the candidate region image with binary image to get Zebra-crossing image. Then through the information of stripes’ feature like width, height, width ratio and angle to extract contour information. Due to the width of black and white stripes and the ratio of both kinds of stripes is constant; the Zebra-crossing can be extracted through contour image scanning.

The detailed steps of the algorithm are as follows:

Step 1: Sensor installation

Step 1-1: Install a monocular camera on one rear view mirror of the intelligent car and make the camera parallel with the longitudinal coordinate axis of the car, the relative displacement of abscissa is zero.

Step 1-2: Install a GPS device on the roof the intelligent car and make sure it’s located on the center line of the intelligent vehicle;

Step 2: Receive real time information from intersection trigger points, where the intersection refers to the closest intersection on the road. The information structure of the trigger point is (intersection types, distance), of which the distance refers to the distance between intelligent car and the intersection center in meters.

Step 3: When the receiving distance is less than 100 meters, the zebra crossing recognition program is launched immediately. On the contrary, the zebra crossing recognition program is closed when the receiving distance is more than 100 meters.

Step 4: Zebra crossing identification;

Step 4-1: Acquire the original pavement image Img from the camera, Firstly, get the gray level image grayImg through grayscale processing the Img. Then applying Inverse Perspective Mapping grayImg is transformed to the aerial image birdeyeImg. Then through the adaptive birdeyeImg processing to obtain the binary image binaryImg.

Step 4-2: Extract candidate regions by bipolar segmentation of the birdeyeImg in the step 4-1, and fuse the candidate regions with the binaryImg. Then extract profiles from the fused image. Finally, according to the constraint conditions the zebra crossing is discerned. The constraint conditions of the zebra crossing refer to width of the write Zebra line $W$ ($40cm < W < 50cm$); height of the write Zebra line $H$ ($300cm < H < 500cm$); the gap between two consecutive strips $w$ ($55cm < w < 65cm$). The number of Zebra lines is between 3 and 10, $3 < N < 10$.

Step 5: Assess the stability of the Zebra-crossing identification results. Here stability refers to correctly identifying the zebra line continuously. Using the continuity of the inter frame correlation, determine the correctness and continuity of the stop and recognition results. Namely, when 3 consecutive frames identify the Zebra crossing, it is considered that the zebra crossing identification is stable.
Figure 10. Experimental Results of the Key Steps

(a) road gray level image; (b) extraction images of candidate regions; (c) road binarization image; (d) fusion image; (e) the results of contour extraction; (f) recognition results

5. Result and Comparison

To evaluate the performance of the proposed method we used 200 images with different backgrounds taken by a monocular camera. Among them, the size of the images is 1000w × 290h pixels. Figure 11, shows some samples of the experimental images.

Figure 11. Different Steps in Detecting the Existence of a Crossing

The complete result of the detection of the existence of crossing is summarized in Table 1. From this table, we see that the proposed algorithm is quite successful in detecting the existence of crossings from real road images. The method has not made any dangerous (false positive) error such that it decides the existence of a crossing for a scene of without crossing.

Table 1. Detection Result Summary

<table>
<thead>
<tr>
<th>Decision</th>
<th>Image with Zebra-crossing</th>
<th>Image without Zebra-crossing</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>315</td>
<td>1017</td>
</tr>
<tr>
<td>False</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>318</td>
<td>1019</td>
</tr>
</tbody>
</table>
The advantages of this method are as follows

(1) Making full use of spatial correlation improves the recognition efficiency of Zebra crossing and reduces the cost and the chance of error identification. It overcomes the disadvantages like expensive and high error detection which caused by the traditional identification methods only based on the video image recognition. That’s because the traditional identification methods need to identify the entire image including the roads without zebra crossing, such as highway road. Thus, computational cost was increased, the cost of hardware was improved and the mistake rate was also increased.

(2) Making full use of the temporal correlation will improve the accuracy of zebra crossing recognition. Temporal correlation is defined as the correlation between frames. It can estimate the stability and correctness of the identification, according to the recognition result which is based on the small changes among the frame images. Although this could cause missing Inspection, the recognition accuracy is greatly improved. And this meets the requirements of intelligent vehicle decision, which is the recognition result returned to the decision layer must be reliable. Otherwise, the judgment will be interfered. It is very unstable for traditional recognition methods to only deal with the single frame image, even the false detection happens quite often, that is, the zebra crossing could be identified when there is no zebra crossing on the road. That does not meet the actual application needs.

References

of the art and

... 2000)

... 2010, pp. 1141–1146.


... “A new approach to linear filtering and prediction problems”, Trans. ASME J. Basic Eng., no. 82, (1960), pp. 35–45.


... “Lane detection & localization for UGV in urban environment”, in Proc. IEEE Int.Conf. ITSC, Qingdao, China, (2014) Oct., pp. 590-596.


