

Nighttime Pedestrian Ranging Algorithm Based on Monocular Vision

Hongxia Wang, Xue-Xue Kang, Yang Yu

*School of Information Science and Engineering, ShenYang Ligong University, ShenYang 110159,
China*

Abstract: *Since the traditional computer vision ranging algorithm is imperfect in pertinence and precision, night time monocular vision pedestrian ranging method is proposed for vehicular infrared night vision goggles. Firstly, the method calibrated the internal and external parameters of infrared night-vision goggles, then, it corrected distortion of collected Vehicular Infrared Night Vision Image, and finally it ranged objective pedestrians by using night time monocular vision pedestrian ranging algorithm. The experimental results show that this method has the characteristics of pertinence, high precision and good real-time, and has good practicability.*

Keywords: *Ranging algorithm, infrared night vision goggles, monocular vision, distortion correction.*

1. Introduction

Vehicular infrared night-vision goggles, based on the principle of infrared thermal imaging, are designed as the auxiliary driving system, they have a wide application prospect and huge commercial value [1, 2]. In the field of infrared night vision image ranging, the internal and external parameters calibration of night-vision goggles occupies a large proportion, the calibration quality directly impacts the ranging precision. Therefore, in the auxiliary driving system, the calibration of goggles in ranging system is the key, while how to select accurate, simple, fast calibration method [3, 4] is one of the key technologies in the field of computer vision.

One of the important tasks of computer vision is to obtain the distance, which is relative to the camera target object in the scene. The more mature model visual range's is geometrical optics method and the binocular stereo vision. Geometrical optics method includes from the focal method and focusing method. Point spread function ranging algorithm based on sensitivity analysis method, which is proposed

by Scheelmer and Kiryati [5], it does not need to find the focus of precise location, but the biggest difficulty is the accurate calibration of the model. Focusing evaluation function based on the improved Laplace (SML) is proposed by N a y a r and N a k a g a w a [6] its principle is simple, but its precision is severely limited owing to the hardware conditions, it needs to focus on the depth of the different areas, so it rarely used. Binocular stereo vision ranging algorithm [7] finds the corresponding points by image matching [8], calculating parallax can get the target depth information in the scene, but its computational complexity is higher, the system is complex, so it isn't suitable for automotive auxiliary driving system research. Because the vehicular infrared night vision system is based on monocular vision of auxiliary driving system, therefore, monocular vision pedestrian ranging algorithm is proposed for automotive auxiliary driving system, relative to other ranging algorithm of binocular stereovision, the method is simple and easy to use, and has a broad application prospect.

2. Calibration of infrared night-vision goggles

2.1. The mutual transformation relationship between coordinate system

The image coordinate system includes the physical image and the pixel image coordinate system. For any image, the rectangular coordinate system O_0UV in the image is defined, and the point of the top left corner is defined as the origin, each coordinate corresponds to a pixel point, the coordinate system is shown in Fig. 1. Since the coordinates of the object cannot change in the image coordinate system and world coordinates directly, the relationship between the world coordinate and image coordinate is established through geometrical relationship. Geometric relationship is built by the camera coordinate system as shown in Fig. 1.

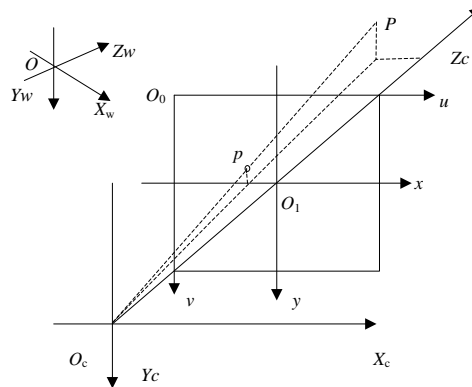


Fig. 1. Schematic diagram of the coordinate system

$O_c X_c Y_c Z_c$ is the camera coordinate system. In Fig. 1, X_c axis and the image coordinate system of the u axis parallel to each other, Y_c axis and v axis parallel to each other, Z_c axis perpendicular to the image plane O_0UV , and is the optical axis of the camera.

The camera focal f is the distance of the optical center along the main shaft to the image coordinate system, $OX_wY_wZ_w$ that is selected in space is the world coordinate system, it is used to describe the location of the camera and the object. In the image physical coordinate system, the point P in the camera coordinate system can be represented as

$$(1) \quad \begin{cases} X = f_x / z, \\ Y = f_y / z. \end{cases}$$

The type converts to the image coordinate relations:

$$(2) \quad \begin{cases} u - u_0 = \frac{X}{d_x} = s_x X, \\ v - v_0 = \frac{Y}{d_y} = s_y Y. \end{cases}$$

In (2), $s_x = 1/d_x$, $s_y = 1/d_y$, its physical meaning is the pixel number per unit length in the X axis and Y axis direction. Transformation relationships between object point P and image point P in the image pixel coordinate system are as follows:

$$(3) \quad \begin{cases} u - u_0 = \frac{fs_x x}{z} = \frac{f_x x}{z}, \\ v - v_0 = \frac{fs_y y}{z} = \frac{f_y y}{z}. \end{cases}$$

In (3), $f_x = fs_x$, $f_y = fs_y$, the equivalent focal length is Defined as f_x , the equivalent focal length in the Y direction is defined as f_y .

2.2. The infrared night-vision goggles parameter calibration

In order to determine the inside and outside parameters of the camera lens, the camera needs to be calibrated. Camera self-calibration method does not rely on calibration reference, it is mainly used for relative accuracy, the calibration method is more flexible, but the stability and reliability is less than that of the traditional camera calibration method. Usually, the traditional camera calibration method applies to the parameters of the camera without changing, and the occasion of relative accuracy.

The two-step calibration method is based on radial calibration constraints is proposed by Roger Tsai [9, 10], its first step solves the camera external parameters using the least square method. The second step is to determine whether the camera has no lens distortion, if without distortion, then the camera internal parameters are solved by a linear equation, otherwise, the camera internal parameters value contains the distortion factor. This process is to determine the camera internal and external parameters, the translation vector t , the rotation matrix R and distortion factor k , the distortion factor is the focal length without radial distortion factor.

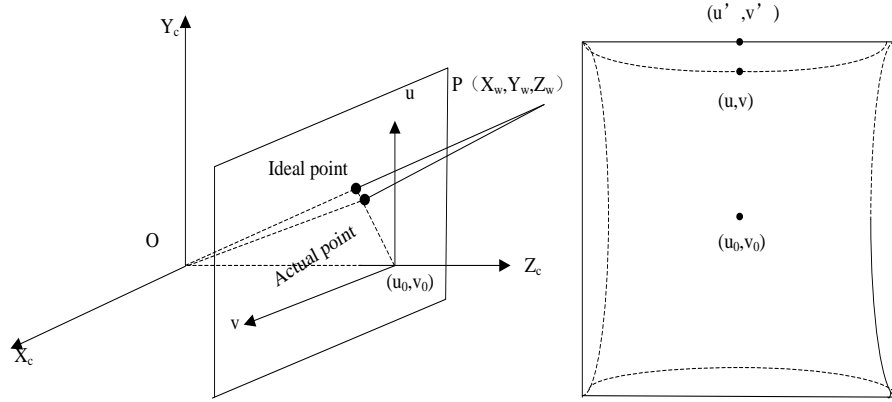


Fig. 2. The ideal position and the actual position of before and after radial distortion

Homogeneous coordinates of any point P in world coordinate system is $(X_w, Y_w, Z_w)T$, the corresponding coordinate in the camera coordinate system is $(X_c, Y_c, Z_c)T$, because of distortion factor k , then the corresponding function relations is as follows:

$$(4) \quad \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} R & t \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}.$$

In the method, which was proposed by Tsai, the matrix K was

$$(5) \quad K = \begin{bmatrix} fs & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Camera calibration method based on two-dimensional planar target is non-linear, as proposed by Zhang Zhengyou [11] (see also [12]). The image is collected through different perspectives and different angles, and five inside parameters of the camera are calibrated.

Any one of the feature points on two-dimensional planar target writes for (x_{wi}, y_{wi}, z_{wi}) , its image coordinates marks (u_i, v_i) , the corresponding functional relationship is as follows:

$$(6) \quad s \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = M \begin{bmatrix} n & o & a & p \end{bmatrix} \begin{bmatrix} x_{wi} \\ y_{wi} \\ 0 \\ 1 \end{bmatrix} = M \begin{bmatrix} n & o & p \end{bmatrix} \begin{bmatrix} x_{wi} \\ y_{wi} \\ 1 \end{bmatrix},$$

where M is the camera intrinsic matrix parameter, s is the coefficient for depth, H is the single rigid matrix, which is mapped from the world coordinate system to the image space,

$$(7) \quad H = M \begin{bmatrix} n & o & p \end{bmatrix} = [h_1 \quad h_2 \quad h_3].$$

The camera distortion is represented using the lens distortion of fourth order radial distortion model:

$$(8) \quad M = \begin{bmatrix} k_x & k_s & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In order to solve the linear camera internal and external parameters, can make:

$$(9) \quad B = M^{-T} M^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix}.$$

In (9), B is a symmetric matrix, camera parameter expressions are:

$$(10) \quad \begin{cases} v_0 = (B_{12}B_{13} - B_{11}B_{23}) / (B_{11}B_{22} - B_{12}^2), \\ k_x = \sqrt{c / B_{11}}, \\ k_y = \sqrt{cB_{11} / (B_{11}B_{22} - B_{12}^2)}, \\ k_s = -B_{12}k_x^2k_y / c, \\ u_0 = k_s v_0 / k_y - B_{13}k_x^2 / c, \\ c = B_{33} - [B_{13}^2 + v_0(B_{12}B_{13} - B_{11}B_{23})] / B_{11}^2. \end{cases}$$

The camera internal parameters are determined, through single rigid matrix:

$$(11) \quad H = \lambda M [n \quad o \quad p] = [h_1 \quad h_2 \quad h_3].$$

In (11), λ is a constant factor, finally the camera external parameters are got as follows:

$$(12) \quad \begin{cases} \lambda = I / \|M^{-1}h_1\| = I / \|M^{-1}h_2\|, \\ n = \lambda M^{-1}h_1, \\ o = \lambda M^{-1}h_2, \\ a = n \times o, \\ p = \lambda M^{-1}h_3. \end{cases}$$

2.3. The infrared night vision image distortion correction

2.3.1. Distortion correction algorithm

Assume that each axis has the same distortion coefficient, distortion coefficient is M , the conversion formula is

$$(13) \quad \begin{cases} x = x' + x' \left[k_1(x'^2 + y'^2 + k_2(x'^2 + y'^2)^2) \right], \\ y = y' + y' \left[k_1(x'^2 + y'^2 + k_2(x'^2 + y'^2)^2) \right]. \end{cases}$$

In (13), (x', y') is the ideal coordinates of orthoscopic imaging point, (x, y) is the corresponding actual coordinates, equations are got from the parameter model:

$$(14) \quad \begin{cases} u = u_0 + k_x x + k_y y, \\ v = v_0 + k_y y. \end{cases}$$

Ignore the k_s and getting into (14) can obtain

$$(15) \quad \begin{cases} u = u' + (u' - u_0) \left[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2 \right], \\ v = v' + (v' - v_0) \left[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2 \right]. \end{cases}$$

In (15), the image of the main point coordinate is (u_0, v_0) , the actual image coordinate is (u, v) , without the distortion of the ideal image coordinate is (u', v') .

On the basis of inside and outside the parameters of the camera, the infrared night vision image was ready for distortion correction. About 20 infrared night-vision goggles images were demarcated: first of all, the distortion coefficient was 0, it was no distortion, the ideal image coordinate (u', v') was calculated. Second, imaging feature points corresponding to the imaging plane of ideal image coordinates (x', y') was computed by the camera outside the parameter equation. Then, the distortion coefficient k_1 and k_2 were solved using least square method. Finally, the camera interior and exterior parameter was optimized by (15) and distortion coefficient, and the actual space position could be ensured by multiplying the proportion coefficient.

In fact, the image distortion correction was to a certain amount of rotating camera coordinate system, the process would be back projection infrared image, so the projection matrix was the real target of the infrared image distortion correction, the projection matrix was obtained by calibrating infrared night-vision goggles. At the same time, the relationship between the scenery of the world coordinates and the corresponding image coordinates was determined by the projection matrix, so the image distortion correction could be completed using The projection matrix transformation.

2.3.2. Distortion correction results

Matlab provides the calibration toolkit of the lens and image, selecting mode interface was shown as Fig. 3, calibration toolkit interface was shown as Fig. 4. The first step was to select the Standard pattern in Fig. 3 calibration tool into interface shown in Fig. 4, the second step was to click on the image names button to calibrate the image selection, the third step was to click on the image Read button to calibrate the image read, The fourth step was to click on Extract grid corners button of image Extract to calibrate and extract the Angle point of the image, and the fifth step was to click on the calibration button to start calibration. After the fifth step was completed, the inside parameters of infrared night-vision goggles lens completed calibration, and clicking on the Save button to Save the results of calibration. Finally, clicking on the Undistort image button, the original image distortion correction was completed, it used the saved goggles parameters of the last step.



Fig. 3. Calibration kit mode selection interface



Fig. 4. Calibration kit interface

An infrared night vision image was collected to do distortion correction analysis in fog weather during the night, uncorrected image and the image distortion correction results were shown in Fig. 5.

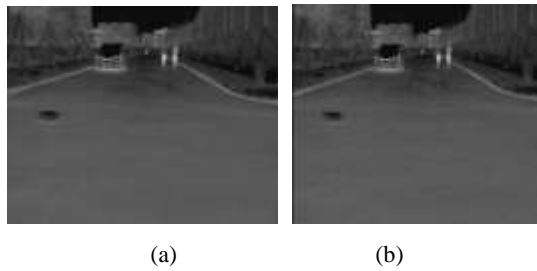


Fig. 5. The image of before and after the correction: Before correction (a); after correction (b)

It can be seen that the image has certain distortion without calibration from the experimental results, the distortion causes a certain impact the late experimental error will reduce using the image with calibration distortion. In order to improve the accuracy of the late experiment, it is necessary to deal with the image distortion correction. At the same time, before correction, the lens distortion coefficient of the infrared night-vision goggles is needed to calculate for image distortion correction used.

3. Monocular vision of pedestrian ranging algorithm implementation

Vehicular infrared night-vision goggles collected two frames image in the corresponding time interval, target pedestrian imaging was respectively in the two frames, assuming that the car's speed was greater than the pedestrian walking speed, pedestrian moving distance was neglected in the image, so the pedestrian imaging was in the same position of the scene. The camera coordinate system of 3D coordinate of pedestrians was concluded by image processing and monocular vision

ranging algorithm. Pedestrian ranging algorithm based on monocular vision was divided into four steps.

First of all, the car carrying infrared night-vision goggles was going at a constant speed along the road, infrared night-vision goggles calibration had been finished, the car went into the distance scene, the schematic diagram was shown in Fig. 6.

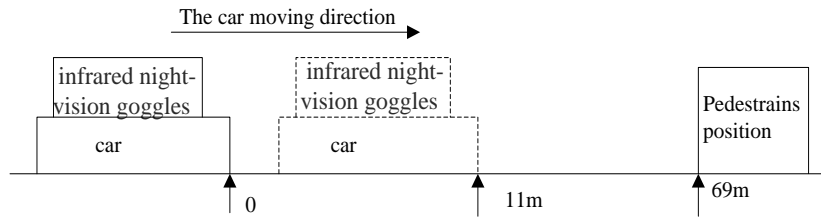


Fig. 6. Schematic diagram of the car

Second, the car along the same direction, the goal pedestrians in front of the car was captured in a certain time interval. Assuming the first image that was collected by goggles was the vista, the second image was the close shot figure. 3D scene is shown in Fig. 7.

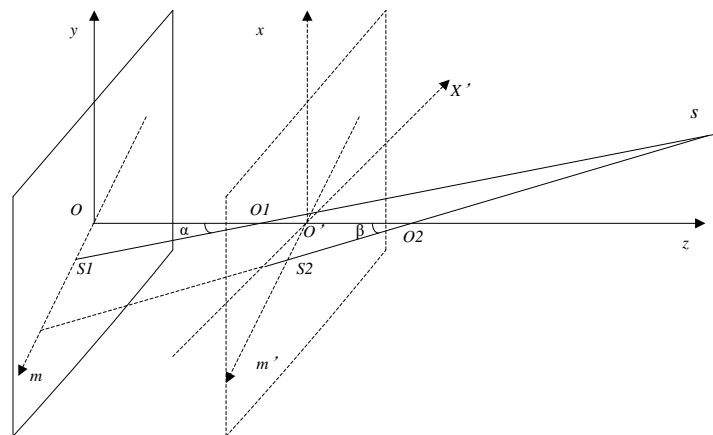


Fig. 7. The two consecutive frames image coordinates after moving

Finally, the pedestrian feature points was calculated in the image of the 3D coordinates, it was shown in Fig. 8. The space rectangular coordinate system $Oxyz$ of infrared night-vision goggles was established, the origin O was the optical center location before moving, before and after the two images was the image plane after imaging. As shown in Fig. 8, the car was going at a constant speed along a straight line, imaging before and after the moving distance was d , S was at any point within the scene, S_1 and S_2 was the imaging before and after moving, the scenario interior point S to the first frame image optical center distance was the OS . The fixed focal length of night-vision goggles was f , and $OO_1 = OO_2 = f$, a plane SOM of the z axis was perpendicular to the plane xOy .

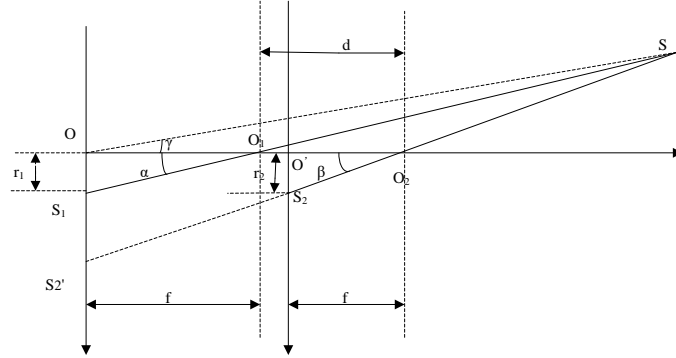


Fig. 8. Coordinate of target point in the xoy plane

The imaging plane is analyzed separately as shown in Fig. 8. The image point S_1 and S_2 to the optical center distance is respectively r_1 and r_2 . In the imaging before and after the two frames, Fig. 9a of the space is larger, and the Fig. 9b has more detail information. In the Fig. 9, $O_1O_2 = d$ is known, $\angle\alpha$ and $\angle\beta$ can be determined, so the triangle SO_1O_2 can be solved.

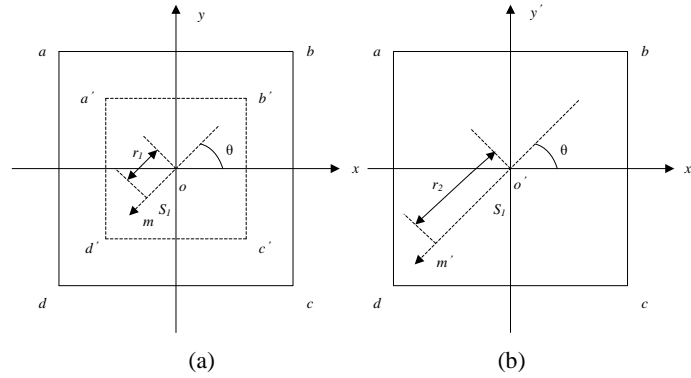


Fig. 9. The plane imaging of moving car before and after

In the camera coordinate system $Oxyz$, assuming that the coordinates of the point S are (x, y, z) , the coordinates of the point S_1 are (i, j) in the plane xOy , the coordinates of the corresponding point S_2 are (k, n) in the plane $x'O'y'$, the coordinates of the point S in the plane $Oxyz$ are derived by trigonometric function and geometric relationships:

$$(16) \quad \begin{cases} x = -jd \frac{\sqrt{k^2+n^2}/f + 2f \left[1 - f^2 / (f^2 + i^2 + j^2) \right] / \sqrt{i^2 + j^2}}{\sqrt{k^2+n^2} + \sqrt{i^2 + j^2}}, \\ y = -id \frac{\sqrt{k^2+n^2}/f + 2f \left[1 - f^2 / (f^2 + i^2 + j^2) \right] / \sqrt{i^2 + j^2}}{\sqrt{k^2+n^2} + \sqrt{i^2 + j^2}}, \\ z = f + d + d \frac{2f / (f^2 + i^2 + j^2) - \sqrt{i^2 + j^2}}{\sqrt{k^2+n^2} + \sqrt{i^2 + j^2}}. \end{cases}$$

In (16), the corresponding points S_1 and S_2 coordinates of the target pedestrians are respectively (i, j) and (k, n) in two frames image, the focal length of the goggles is f , the speed of the car is v , the time interval of two consecutive frames is Δt , auto mobile distance is $d = v \Delta t$, the camera coordinate system and world coordinate system overlap, so the measured 3D coordinate is the absolute coordinates of the world coordinate system. In the Fig. 6, $OO_1 = d$, $zO = z$, the distance between cars and pedestrians is $L = zO_1$, and $zO_1 = zO - OO_1$, the monocular vision of pedestrian ranging formula can be derived, which is

$$(17) \quad L = z - d = f + v \cdot \Delta t \cdot \frac{2f / (f^2 + i^2 + j^2) - \sqrt{i^2 + j^2}}{\sqrt{k^2 + n^2} + \sqrt{i^2 + j^2}}.$$

4. Infrared night-vision goggles ranging results analysis

The gold IR31 series model is adopted by infrared night-vision goggles, the effective focal length is 14.8 mm/F0.96, viewing Angle is $36^\circ \times 27^\circ$, the pixels is 720×576 , night-vision goggles of internal and external parameters calibration has been finished. The physical diagram and the installation location of infrared night-vision goggles in the car is shown in Fig. 10.

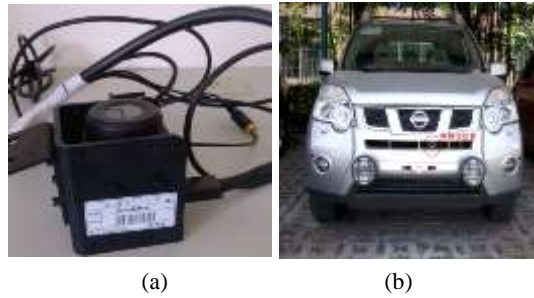


Fig. 10. Physical map and installation position of night vision goggles: Physical map (a); installation position on the car (b)

Infrared night-vision goggles without calibration collected a set of images, as shown in Fig. 11, the green rectangle is the recognition of the pedestrian.

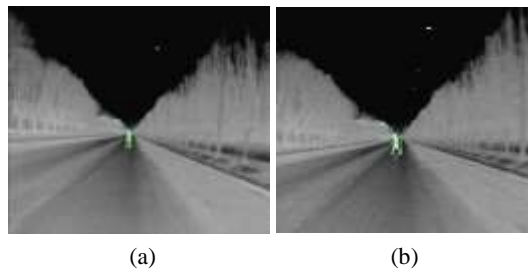


Fig. 11. Uncorrected infrared night vision image: First frame image (a); second frame image (b)

Due to the infrared night-vision goggles lens has distortion, the space position of the image reaction is not accurate. After night-vision goggles of internal parameters calibration has been finished, gathering a group of images is shown in Fig. 12.

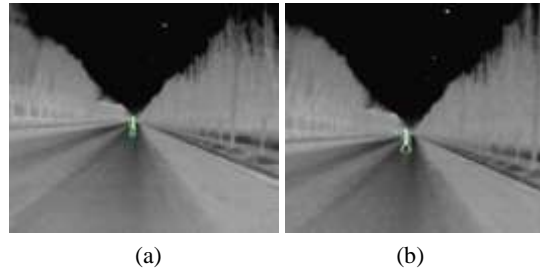


Fig. 12. Corrected infrared night vision image: First frame image (a); second frame image (b)

The world coordinate of the pedestrians in the correction of image is calculated with monocular vision of pedestrian ranging algorithm, and the calculated value and the actual distance of the car are compared, the relative error is calculated. the center of gravity said to identification of pedestrian position, the barycentric coordinates of pedestrians rectangular is calculated by computer, the car's speed is 60 km/h, the time interval is 1 s. The distance of pedestrians and vehicles is calculated about $L = 69.92$ m using ranging algorithm, and the actual distance of people and the vehicle's is $L_{\text{real}} = 68.26$ m, the relative error is about 1.96%, it is in a reasonable range. Ten images is collected randomly, and its sample and actual value is calculated as shown in Table 1.

Table 1. The calculation and actual value of monocular vision

Calculated value, m	Actual value, m	Relative error, %
130.76	134.93	3.09
120.70	123.38	2.17
109.98	112.13	1.92
99.72	101.56	1.81
88.63	90.34	1.89
77.79	79.39	2.02
66.92	68.26	1.96
56.09	57.08	1.74
44.51	45.27	1.68
33.90	34.45	1.61

It can be seen from Table 1 that within the scope of the distance of 110 m, the relative error is within 2%, as the distance of the vehicle and pedestrian became close, the length of the imaging in the goggles becomes big, its center of gravity values also become relatively accurate, so the distance of cars and pedestrians is shorter, error became smaller. When the distance measuring value is less than the

safety threshold set, the car warned drivers in order to prevent the possibility of traffic accidents.

5. Conclusion

As the effect of the traditional pedestrian ranging algorithm is not good, the nighttime pedestrian ranging method based on monocular vision is proposed. The key technology in this method is researched and analyzed, where the experimental results verified that the nighttime pedestrian ranging method based on monocular vision is feasible, it is highly targeted, high precision, and it can meet the requirements of the auxiliary driving system real-time performance.

Acknowledgments: This research was partially funded by the National Natural Science Foundation of China (Grant No 61301256), Supported by Program for Liaoning Excellent Talents in University (No LJQ2013021).

Reference

1. Z h u, X., F. L i u. Digital Image Processing and Image Communication. BeiJing, Beijing University of Posts and Telecommunications Publishing House, 2002, pp. 68-72.
2. Q i d w a i, U. Infrared Image Enhancement Using H_{∞} Bounds for Surveillance Applications. – IEEE Transactions on Image Processing, Vol. **17**, 2008, No 8, pp. 1274-1282.
3. Z h u, S. J., Y. Z h a o. Research of Binocular Camera Self-Calibration Method. – Photoelectric Technology Application, Vol. **28**, 2013, No 3, pp. 36-39.
4. H o u, J., E. X u. A High-Accuracy Calibration Technique of Stereo Vision for Mobile Robot. – Computer Technology and Development, Vol. **24**, 2014, No 2, pp. 92-95.
5. S u b b a r o, M., G. S u r y a. Depth from Defocus: A Spatial Approach. – International Journal of Computer Vision. Vol. **13**, 1994, No 3, pp. 271-294.
6. N a y a r, S. K., Y. N a k a g a w a. Shape from Focus System. – IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. **16**, 1994, No 8, pp. 824-831.
7. L i, H. B., W. D a n. Research of Error-Model on Two Eyes Stereoscopic Measurement System. – Optical Technique, Vol. **32**, 2006, No 1, pp. 24-30.
8. C h e n g, D., Y. L i. Research for Image Matching Based on Improved SIFT Algorithm. – Electronic Design Engineering, Vol. **28**, 2011, No 7, pp. 285-289.
9. T s a i, R. Y. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. – In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, 1986, pp. 364-374.
10. T s a i, R. Y. A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses. – IEEE Transactions on Robotics and Automation, Vol. **3**, 1987, No 4, pp. 323-344.
11. Z h a n g, Z. A Flexible New Technique for Camera Calibration. – IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. **22**, 2000, No 11, pp. 1330-1334.
12. Y u e, X., H. Q i. Calibration of Binocular Stereo-Visoin System Based on Zhang Zhengyou Plane Template Method. – Mechanical Engineer, 2014, pp. 1-2.