

PAPER • OPEN ACCESS

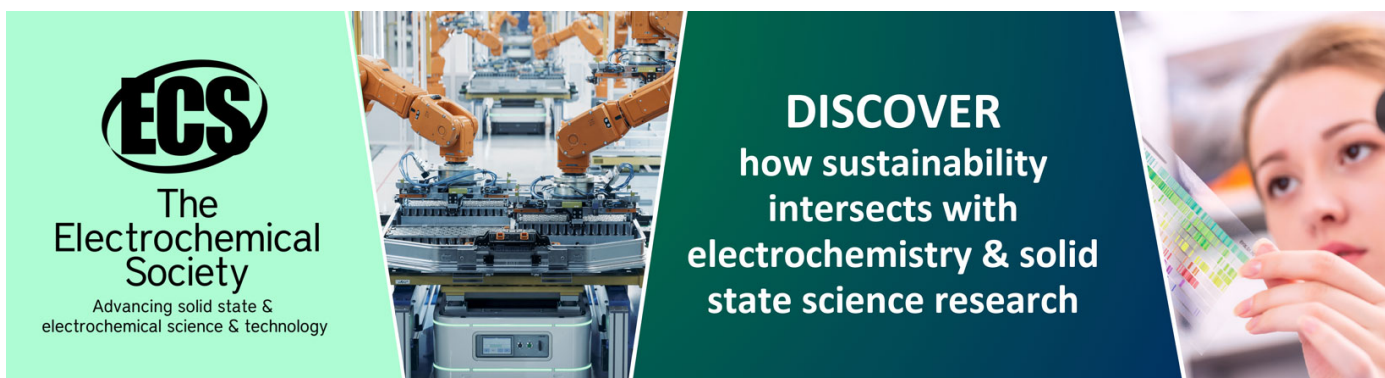
A Rapid Method of Monocular Image Measurement Based on Rectangle Information

To cite this article: Zhuolin Ping and Zhenjun Yue 2020 *J. Phys.: Conf. Ser.* **1631** 012101

View the [article online](#) for updates and enhancements.

You may also like

- [Monocular 3D Object Detection Using Depth Fusion](#)
Dewen Qiao and Hongxia Niu
- [A Survey on Monocular 3D Object Detection Algorithms Based on Deep Learning](#)
Junhui Wu, Dong Yin, Jie Chen et al.
- [A measurement system based on internal cooperation of cameras in binocular vision](#)
Yaqin Zhou, Qingwu Li, Lulu Chu et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

A Rapid Method of Monocular Image Measurement Based on Rectangle Information

Zhuolin Ping* and Zhenjun Yue

College of Communication Engineering, Army Engineering University of PLA, 1
Haifu lane, Guanghua Road, Qinhuai District, Nanjing, China
Email: jolinping@163.com

Abstract. A method of rapid measurement of monocular image based on rectangle information is proposed in this paper. Based on the pinhole imaging model, a fast method of calculating the object size in monocular image is derived by fitting the straight line and then finding the intersection point of the straight line. The algorithm can not only measure the size and absolute depth of the object in monocular image, but also meet the conditions of the specific pinhole imaging model with a distance far greater than the focal length. The complexity of the algorithm is low, and it can meet the requirements of real-time calculation. It provides a new idea for the use of small and micro-computing platform.

1. Introduction

With the development of computer vision technology, there is a strong application demand in the field of vehicle monocular ranging model based on the principle of monocular vision and the corresponding ranging algorithm [1], mobile robot positioning, navigation system [2], and UAV positioning [3]. To complete 3D scene modeling, it is very important to measure the depth, relative distance and other information of 3D scene based on a single image, that is, monocular vision ranging.

Different from binocular vision, monocular vision ranging can only measure distance information by previous knowledge, such as image size change [4], which is very difficult. Some of the existing methods mainly include the indirect measurement of single strain matrix [5] and the direct measurement based on invariants [6]. The calculation of single strain matrix method is complex, and the applicability of direct measurement method based on invariants is narrow. However, with the introduction of neural networks and deep learning methods, more and more scholars begin to study in the field of monocular vision ranging. Typical research results include: Lim J H et al. proposed a hybrid positioning method combining single frequency GPS receiver and monocular vision sensor [7]; Wu et al. studied the integration of monocular camera and laser rangefinder [8]; Han, Lessmann, et al. proposed a monocular vision distance estimation algorithm based on probability [9-10]; Kendall et al. used the deep convolution neural network to explore the method of positioning based on geometric scene learning error cameras [11] and the monocular vision image measurement method based on entropy and weighted Hu's invariant moment proposed by He Lixin of the University of Science and Technology of China [12]. Compared with the monocular vision ranging method based on neural network and deep learning, which needs a lot of data for training and learning, the model-based vision location method is more available. Therefore, a rapid solution algorithm based on the distance information of edge detection and Hough transformation is proposed. Experiments show that the error and time complexity of the algorithm meet the practical requirements.



2. Camera Shooting Process Modeling

The reliability study of linear model of monocular camera by Jia and others [13] shows that when the lens object distance is much larger than the focal length, the lens model can be approximately transformed into a pinhole imaging model, and can obtain high accuracy. Based on the pinhole imaging model, this paper deduces a rapid method to calculate the size of objects in monocular image.

2.1. Space Calibration

As shown in figure 1, point O is the center of the camera lens, $\hat{O}$ is the center of the image plane, make O is the coordinate origin, $O\hat{O}$ direction is z axis direction, make $|O\hat{O}| = f$, s is the ratio coefficient between the actual size of the camera sensitive device and the pixel coordinates of the image, so that the main distance of the image is $u = s \cdot f$. The image of rectangle $A'B'C'D'$ on image plane is quadrilateral $ABCD$, A point coordinate is (x_a, y_a, u) , B point coordinate is (x_b, y_b, u) , C point coordinate is (x_c, y_c, u) , D point coordinate is (x_d, y_d, u) . $P_i'(x, y, z)$ is an end point of the line segment to be measured in the real scene. $P_i(x_p, y_p, u)$ is the corresponding point in the image and i represents the serial number of the end point.

Suppose vp_1, vp_2 are the coordinates of vanishing points formed by two sets of orthogonal parallel lines.

It is easy to derive:

$$u = -\sqrt{vp_1 \bullet vp_2} \quad (1)$$

2.2. Model Establishment

For the convenience of derivation, the spatial coordinates of figure 1 are simplified to figure 2. In figure 2, the point O is the camera lens center, the rectangular reference $A'B'C'D'$ imaging on the plane is a quadrangle $ABCD$, M' is the intersection of segment $A'C'$ and $B'D'$, M is the intersection of segment AC and BD .

For ease of analysis, assuming that the point A coincides with A' , the simplified map of the mapping relationship between the like plane and the actual plane is shown in figure 3. Set $\angle AOM = \alpha$, $\angle AMO = \gamma$, $\angle A'M'O = \gamma'$, $\angle MOC = \beta$. According to the Sine theorem:

$$\begin{cases} \frac{AM}{\sin \alpha} = \frac{AO}{\sin \gamma} \\ \frac{CM}{\sin \beta} = \frac{CO}{\sin \gamma} \end{cases} \Rightarrow \frac{AM}{CM} \bullet \frac{\sin \beta}{\sin \alpha} = \frac{AO}{CO} \quad (2)$$

$$\begin{cases} \frac{A'M'}{\sin \alpha} = \frac{A'O}{\sin \gamma'} \\ \frac{C'M'}{\sin \beta} = \frac{C'O}{\sin \gamma'} \end{cases} \Rightarrow \frac{A'M'}{C'M'} \bullet \frac{\sin \beta}{\sin \alpha} = \frac{A'O}{C'O} \Rightarrow \frac{\sin \beta}{\sin \alpha} = \frac{A'O}{C'O} \quad (3)$$

In the same way, if the point B coincides with B' , we can get:

$$\frac{BM}{DM} = \frac{B'O}{BO} \quad (4)$$

We set

$$\lambda_1 = \frac{AM}{CM} = \frac{C'O}{CO}$$

$$\lambda_2 = \frac{BM}{DM} = \frac{B'O}{BO}$$
(5)

Among them, the values of line segments AM , BM , CM , and DM can be read out directly on the image plane, so that we can obtain λ_1, λ_2 .

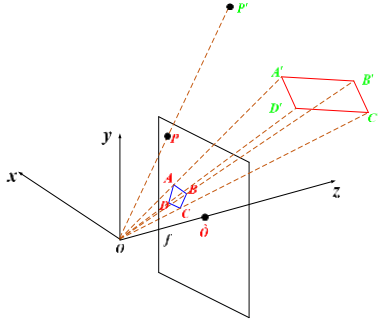


Figure 1. Space calibration.

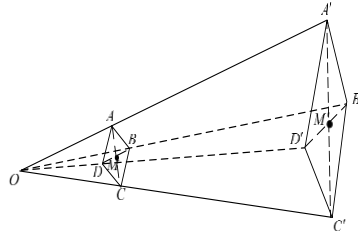


Figure 2. Mapping relationship between image plane and actual plane.

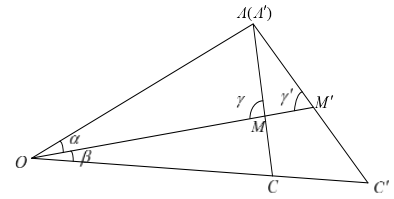


Figure 3. Simplified map of image plane and actual plane.

In fact, point A and A' do not coincide, point B and B' do not coincide either, so that:

$$\overrightarrow{OA'} = \lambda \overrightarrow{OA}$$

$$\overrightarrow{OB'} = \lambda' \overrightarrow{OB}$$
(6)

Then:

$$\overrightarrow{OC'} = \lambda \cdot \lambda_1 \cdot \overrightarrow{OC}$$

$$\overrightarrow{OD'} = \lambda' \cdot \lambda_2 \cdot \overrightarrow{OD}$$
(7)

From equations (6)-(7):

$$\begin{cases} \overrightarrow{A'C'} = \overrightarrow{OC'} - \overrightarrow{OA'} = \lambda \cdot (\lambda_1 \cdot \overrightarrow{OC} - \overrightarrow{OA}) \\ \overrightarrow{B'D'} = \overrightarrow{OD'} - \overrightarrow{OB'} = \lambda' \cdot (\lambda_2 \cdot \overrightarrow{OD} - \overrightarrow{OB}) \end{cases}$$
(8)

In rectangular reference $A'B'C'D'$, let the diagonal length be L , with the following relationship:

$$|\overrightarrow{A'C'}| = |\overrightarrow{B'D'}| = L$$
(9)

It can be concluded from (8) and (9) that:

$$\begin{cases} \lambda = L / |\lambda_1 \cdot \overrightarrow{OC} - \overrightarrow{OA}| \\ \lambda' = L / |\lambda_2 \cdot \overrightarrow{OD} - \overrightarrow{OB}| \end{cases}$$
(10)

Union (6)-(7) can calculate the coordinates of the rectangular reference in the real world. Furthermore, the normal vector of the plane of the reference rectangle is:

$$\vec{n}_1 = \overrightarrow{A'C'} \times \overrightarrow{B'D'}$$
(11)

2.3. Derivation of Distance Formula

The key to solve the length of line segment is to find the coordinates of the end point. If the end point of the line segment and the rectangular reference $A'B'C'D'$ is coplanar and the coordinates and normal vectors of a point on the plane are known, there are:

$$\overrightarrow{OP'_i} = \overrightarrow{OA'} \bullet \vec{n}_1 / (\overrightarrow{OP_i} \bullet \vec{n}_1) * \overrightarrow{OP_i} \quad (12)$$

The absolute depth from this point to the center of the lens is:

$$\text{Depth} = |\overrightarrow{OP'_i}| \quad (13)$$

The distance between the point i and the point j is:

$$d_{ij} = |\overrightarrow{P_i P'_j}| \quad (14)$$

If the line segment to be determined and the rectangular reference $A'B'C'D'$ is not coplanar, the solution can be completed according to the above formula after the normal vector of the plane where the line segment is located has been solved according to the spatial geometry knowledge and the plane has been calibrated.

3. Intelligent Line Detection Method

In order to reduce the pixel pick-up error, when calculating the image vanishing point and obtaining the reference rectangle vertex, the intersection point of the line can be obtained by fitting the line. The common methods of obtaining lines intelligently include: line detection by least square method [14], minimum distance method, Radon transform line detection and Hough transform line detection [15]. In this paper, Hough transform is used to realize line detection. The steps are as follows:

(1) Image pre-processing. Because the original image is usually large in size and contains many lines, which is not conducive to the selection of target lines, the original image needs to be cut. Take the detection of rectangular reference as an example, the image before and after cutting is shown in figures 4a-4b. At the same time, the image needs to be grayed, as shown in figure 4c.

(2) Edge detection. The purpose of edge detection is to identify the points with obvious brightness change in digital image. The commonly used edge detection operators are Roberts, Sobel, Prewitt, Laplacian, Log/Marr, Canny, Kirsch, Nevitia, and so on. Sobel operator has a good effect on image processing with more gray gradients and noise, and its edge location is more accurate, so it has been widely used. In this paper, Sobel operator is used to process image, as shown in figure 4d.

(3) Hough line detection. The line in the image space is transformed to the point in the parameter space, and the line is detected by statistical characteristics. As shown in figure 5, the points (x_1, y_1) , (x_2, y_2) and (x_p, y_p) represent the pixel points on the same line in an image. Then these pixel coordinate values can be represented by parameters as follows:

$$\rho = x \cos \theta + y \sin \theta \quad (15)$$

All points on the same straight line must intersect at one point under the parametric equation. At this point, the problem of finding collinear points is transformed into the problem of solving all curve intersections. It is only needed to transform all the pixel points (coordinate values) in the image into the curve in parameter space, and detect the curve intersection point in parameter space to determine the straight line. Figure 4e is the Hough space curve of the image and figure 4f is the detected line.

(4) Restore the line detected in the area map to the original map. Only a simple coordinate transformation is required to display in the original image, as shown in figure 4g.

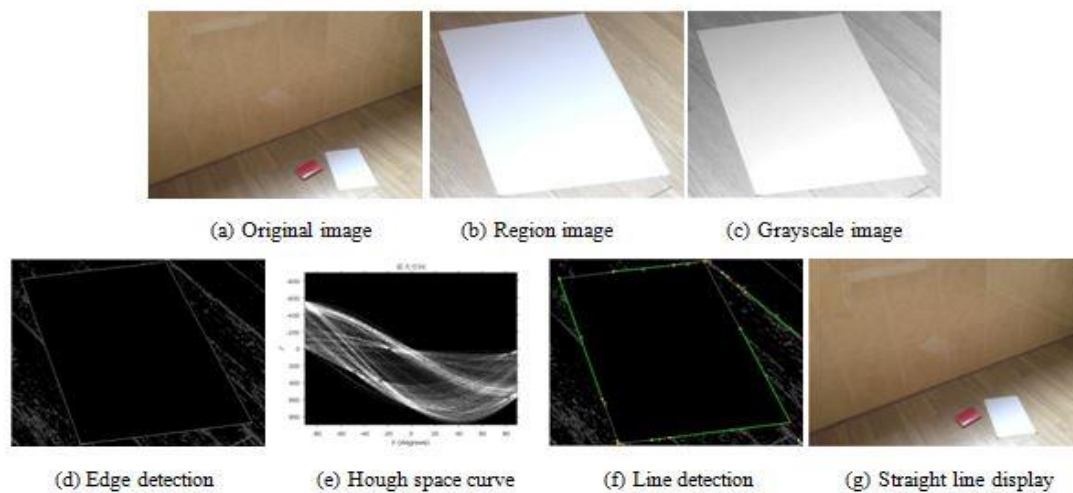


Figure 4. Hough transformation process schematic.

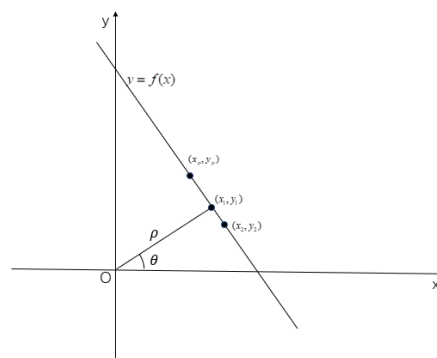


Figure 5. Hough line detection principle.

4. Experiment

4.1. Measurement Accuracy

Define the relative error of the measurement:

$$\text{relative error} = \frac{\text{calculated dimensions} - \text{measured dimensions}}{\text{measured dimensions}} * 100\%$$

The experimental scene shown in figure 6 is designed.



Figure 6. Monocular image to be tested

Figure 6 is an image taken by the mobile phone, in which the size of A4 paper is $297 \times 210 \text{ mm}$. The objects to be measured are the dimensions of floor tiles, mud box and wall tiles, and the absolute depth of the common end point of the line segment. The calculation results retain 2-bit decimal as shown in table 1. For ease of viewing, the measured value is also marked in figure 6, where the value in brackets is the relative errors. The measured value in table 1 is obtained by taking the average value of scale for many times, leaving a decimal place of 1 bit. In order to reduce the influence of other factors on the model error, firstly, clear photos are selected; secondly, the edge of the object to be tested is easy to be extracted; thirdly, in order to reduce the error caused by camera distortion, try to avoid close shot as much as possible, so as to meet the condition of pinhole imaging model with the object distance far greater than the focal length.

It can be seen from table 1 that the maximum error of object size measurement is 5.97%, indicating that the model in this paper can complete the measurement of monocular image. The measurement error of the size of the object on the wall is obviously larger than that of the object on the ground, mainly because the cumulative error is produced when the cross-plane measurement is produced. At the same time, the experiment also shows that this algorithm can measure the absolute depth of the point on the object.

Table 1. Field measurements and measurement errors.

Measure Parameter	AB/cm	BC/cm	DE/cm	EF/cm	GH/cm	HI/cm	B point Absolute depth / cm	E point Absolute depth / cm	H point Absolute depth / cm
Measured Value	40.5	12.5	8.6	13.6	45.1	30.1	145.8	143.6	140.9
Calculated Value	39.15	12.78	8.32	13.19	42.41	28.90	149.27	138.28	132.11
Relative Error	-3.33%	2.27%	-3.32%	-2.99%	-5.97%	-3.99%	2.38%	-3.70%	-6.24%

4.2. Time Complexity

Because of the length of the solution line segment, the key is to find the coordinates of the end point. In order to facilitate the statistics of calculation time, this paper uses the time measurement algorithm to solve the complexity of a single point, the formula is as follows:

$$\text{time complexity} = \frac{\text{total calculation time of } N \text{ points to be solved}}{\text{number } N}$$

The platform used in the experiment is Xiaomi computer pro, and the CPU is Intel ® core™ I5-8250u, 8GB memory, 64bit win10 system, software platform is MATLAB 2017a. The experiment calculates a total of 11 points calculation time (units: seconds). Each point is calculated 10 times to get the average. The curve of the total calculation time with the number of points to be calculated is shown in figure 7.

It can be seen in the figure that the trend of change is approximately linear, and the fitting results are as follows:

$$T = 0.00014316 * n + 0.20049$$

It can be seen from the fitting results that the time for calculating each additional point is:

$$t = 143.2 \mu\text{s}$$

The above analysis shows that the complexity of this algorithm can meet the requirements of real-time calculation.

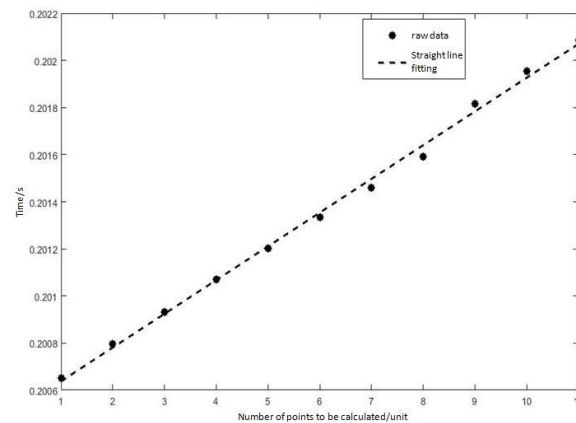


Figure 7. Change curve of total calculation time with the number of points to be calculated.

5. Conclusion

Based on the pinhole imaging model, a fastmonocular image measurement method based on rectangular information is derived in this paper. Compared with He Lixin's measurement of monocular vision image, it does not need to take two images of the same scene, and only a single photo is enough to complete the measurement of the size and absolute depth of the object in the photo. The algorithm complexity in this paper can meet the requirements of real-time calculation. From the experimental results of this paper, the applicability and effectiveness of the algorithm have been confirmed, which provides a new solution for the occasions with limited computing power but the real-time computing requirements. Considering the clarity of the tested image, that is, the quality of the shooting light and the difficulty of the edge extraction of the object to be tested, will have an impact on the accuracy of the algorithm, the next work is to further improve and optimize the algorithm in combination with the above factors.

References

- [1] Zhao H 2019 *Research on Vehicle Borne Monocular Vision Ranging Model and Application* (Inner Mongolia Agricultural University).
- [2] Shou J, Zhang Z, Su Y and Zhong Z 2018 Design and implementation of indoor mobile robot positioning and navigation system based on ROS and lidar *Mechatronics* **36** (11) 76-80.
- [3] Xiong H, Yang R, Liang X, Hou Y and Wang W 2020 AOA 3D positioning method of UAV monocular vision *Computer Engineering and Application* 1-8.
- [4] Xiang Z and Ma B 2020 Research on monocular vision ranging method based on imaging size change *Journal of Armament Engineering* **41** (02) 148-151.
- [5] Peng Q, Yang X, Cheng S and Lv W 2019 Research on the rapid calibration method of line structured light measurement based on the homography matrix *Mechanical and Electrical Engineering* **36** (06) 613-617.
- [6] Skriganov M M 2020 Stolarsky's invariance principle for projective spaces *Journal of Complexity* **56**.
- [7] Lim J H, Choi K H, Cho J, et al. 2017 Integration of GPS and monocular vision for land vehicle navigation in urban area *International Journal of Automotive Technology* **18** (2) 345-356.
- [8] Kai, Wu, Kaichang, et al. 2014 Enhanced monocular visual odometry integrated with laser distance meter for astronaut navigation *Sensors*.
- [9] Han J, Heo O, Park M, et al. 2016 Vehicle distance estimation using a mono-camera for FCW/AEB systems *International Journal of Automotive Technology* **17** (3) 483-491.
- [10] Lessmann S, Meuter M, Muller D, et al. 2016 Probabilistic distance estimation for vehicle tracking application in monocular vision *2016 IEEE Intelligent Vehicles Symposium (IV)*.
- [11] Kendall A and Cipolla R 2017 Geometric loss functions for camera pose regression with deep learning.

- [12] He L 2018 *Research on Depth Measurement Method of Monocular Vision Image* (University of Science and Technology of China).
- [13] Jia X, Yin C and Sui G 2019 Reliability analysis and Research on linear model of monocular camera *Optical Instruments* **41** (01) 60-67.
- [14] Zhao Y, Wu M, Ren Z, Wang Q and Wu K 2018 Line structured light vision technology and hole straightness detection by least square method *Mechanical Design and Research* **34** (01) 141-144.
- [15] Wang Q, Song W and Wang J 2019 Line extraction method and improvement based on Hough transform *Mapping and Spatial Geographic Information* **42** (06) 214-217+221.