# Object Depth Measurement Based on Feature Point Detection

Lai Yufeng
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, China
E-mail: 604612675@qq.com

Wang Zhongsheng
School of Computer Science and Engineering
Xi'an Technological University
Xi'an, China
E-mail: 59483672@qq.com

***Abstract*—In order to simplify the method of obtaining absolute depth information, this paper proposes a method of automatic measurement of object depth information in images by monocular camera without calibration and adjustment of camera parameters, which is used to realize an automatic depth measurement system. This method uses the feature points on the object image to measure the depth of the object to enhance the robustness of the algorithm against partial occlusion or missing of the measured object in the scene. Experimental results show that the method is effective.**

***Keywords-Component; Feature Points; Absolute Depth; Image Segmentation; the Object Matching***

## I. INTRODUCTION

Extracting depth information from two-dimensional plane images is an important research content of stereo vision. Sometimes, it is necessary to extract the depth information of each pixel in two-dimensional images, but more often, it is only necessary to obtain the depth information of the objects of interest. For example, the distance between obstacles and red street lamps in the progress of the robot. The information available for depth extraction includes binocular parallax information, line perspective information and image texture information. In order to avoid the complexity of obtaining depth information by binocular vision, this paper proposes a method of automatically measuring depth information of still objects in images by monocular camera. First, a normal camera is used to shoot the first image at a distance of u, and then all the parameters of the camera are kept unchanged, so that the camera moves along the optical axis for a distance of d to obtain the second image. Secondly, the object in the image is segmented and the feature points in the image are detected. Finally, select the appropriate line segment and substitute it into the derived formula to find the depth u of the object.

## II. PRINCIPLE OF CAMERA IMAGING

The basic principle formula of camera imaging is (1)

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} \tag{1}$$

In the formula, $f$ is the focal length, $u$ is the object distance and $v$ is the image distance.

It is assumed that the object distance at the first image is u and the image height of the object is $h_{image1}$ .The object distance during the second imaging is u plus d, and the imaging height of the object is $h_{image2}$ .As shown in figure 1, since the thin convex lens of the camera is symmetric with respect to the center of the lens, the zoom ratio of all points on the object when they pass through the camera is the same. Assuming that the scaling ratio is k, when the parameters used by the camera remain unchanged during the two imaging, the image height of the two imaging exists as follows:

$$h_{image2} = kh_{image1} \tag{2}$$

Where, $k > 0$ and $k \in R$ , R is real Numbers

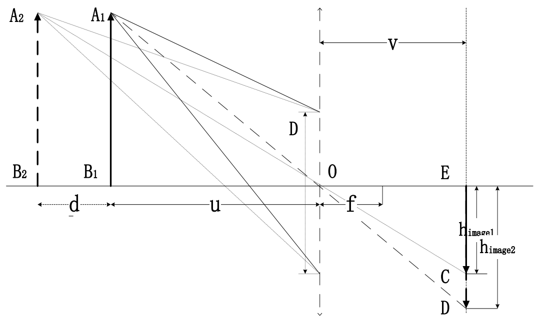



Figure 1. Two lens imaging

As shown in the figure above, it is easy to obtain $\Delta OA_1B_1 \sim \Delta OCE$ and $\Delta OA_2B_2 \sim \Delta ODE$, according to the basic principles of camera imaging and the knowledge of three geometer, it can be known that:

$$\begin{cases} \dfrac{A_1B_1}{B_1O} = \dfrac{h_{image1}}{v} \Rightarrow A_1B_1 = \dfrac{h_1 u}{v} \\ \dfrac{A_2B_2}{B_2O} = \dfrac{h_{image2}}{v} \Rightarrow A_2B_2 = \dfrac{h_2(u+d)}{v} \end{cases}$$
$$\Rightarrow \because A_1B_1 = A_2B_2$$

$$\therefore u = \frac{h_{image2}}{h_{image1} - h_{image2}} d \qquad (3)$$

Substitute equation (2) into equation (3) and get:

$$u = \frac{kh_{image1}}{h_{image1} - kh_{image1}} d = \frac{k}{1-k} d \qquad (4)$$

According to the above equation (4), when the object distance difference d of two images and the zoom ratio of camera imaging are known, the required object depth can be obtained.

## III. IMAGE SEGMENTATION

This paper chooses GrabCut image segmentation method based on graph theory. The basic idea of this method is to treat the original image to be processed as a network graph. As shown in figure 2, s represents the source point, where the node is the pixel point of the object, and t represents the junction point, which represents the pixel point of the background part of the image. The network graph contains two kinds of edges: one is the edge between adjacent pixels in the image, and the size of this edge represents the degree of difference or approximation between two nodes. If the correlation between two nodes is larger, the edge weight will be larger. The other kind of edge is between the pixel point in the image and s and t, and the value is calculated based on the probability that the current point belongs to the target or background.

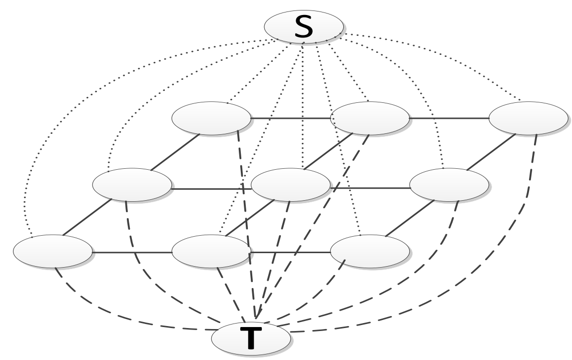



Figure 2. Schematic diagram of image segmentation method

GrabCut needs to manually frame the possible foreground area, and all the areas outside the box are the background area. All foreground region information must be included in the box, otherwise the performance of the algorithm will be affected. Then calculate the distance between all pixels in the image and the target and background marked by the user, and calculate the distance between pixels. The above two kinds of distances are integrated and taken as the energy value of the edge, among which the minimum energy is the optimal solution.

## IV. DETECTION OF FEATURE POINTS AND SELECTION OF STRAIGHT LINE SEGMENTS

After image segmentation, the next step is to find the image of the same object in two different images. That is, the object image matching. There are many research results on this kind of problem, such as using minimum noise separation, manifold learning and discriminating local enhancement matching technology to realize dimension reduction of the data to be processed, so as to carry out image matching quickly and effectively. According to the needs of the problem, this paper proposes a feature point detection method for object recognition and matching.

### *A. Harris-SIFT feature point detection*

In order to obtain more stable feature points and reduce the computational load, this project improves SIFT algorithm based on Harris algorithm from the following aspects (as shown in figure 3).

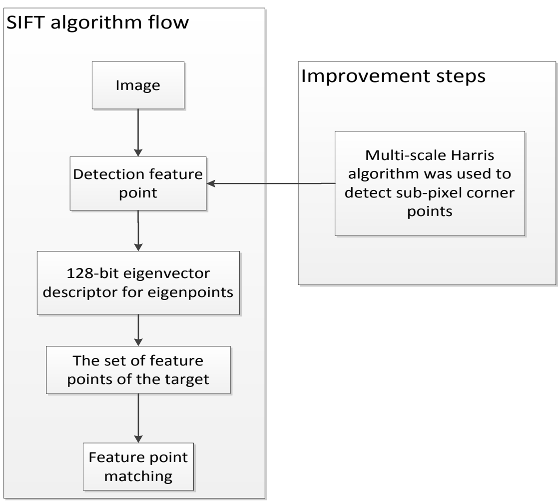



Figure 3. SIFT algorithm improvement steps

#### *1) Harris-SIFT feature point extraction*

Because the distribution of feature points detected by Harris has a distinct geometric significance, the sub-pixel corner points detected by multi-scale Harris are replaced by the feature points detected by SIFT. The improved algorithm can greatly reduce the number of redundant feature points and obtain double stable feature points. This algorithm not only includes Harris algorithm stability to noise interference, but also SIFTS algorithm's characteristics of rotation , brightness and scale are all unchanged. It has better matching effect and wider practicability. The specific flow chart is shown in figure 4.

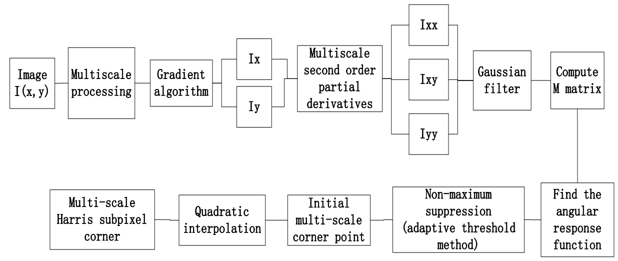
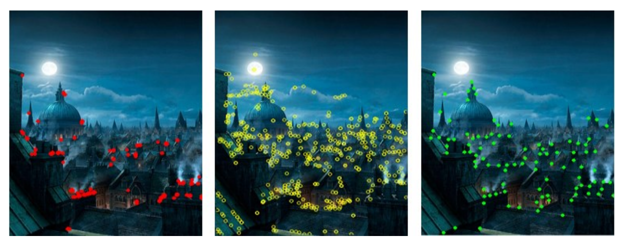


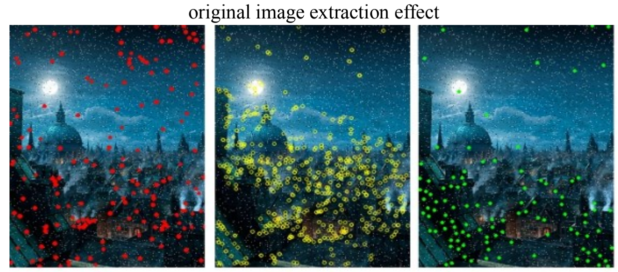

Figure 4. Extraction steps of multi-scale Harris feature points

#### *2) The experimental results*

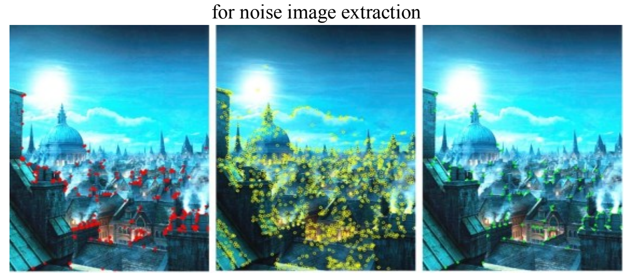
This section compares Harris algorithm, SIFT algorithm and improved Harris-SIFT algorithm. The corner points were extracted from the original image, noise image, light changing image and rotation image. The effect is shown in figure 5. It is concluded that the improved algorithm in this section not only has the stability of Harris algorithm, but also includes the features of SIFT algorithm for care, scaling and rotation invariant.



(a) Harris algorithm, SIFT algorithm, Harris-SIFT algorithm in this paper on the original image extraction effect

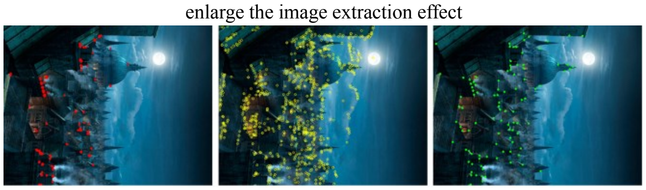


(b) Harris algorithm, SIFT algorithm and Harris-SIFT algorithm in this paper for noise image extraction



(c) the effects of Harris algorithm, SIFT algorithm and Harris-SIFT algorithm on image extraction with light changes

(d) Harris algorithm, SIFT algorithm, Harris-SIFT algorithm in this paper to enlarge the image extraction effect



(e) Harris algorithm, SIFT algorithm, Harris-SIFT algorithm in this paper on the rotation of the image extraction effect

Figure 5. Harris, SIFT and Harris-SIFT effect comparison

### *B. Selection of straight segment by triangle method*

In this paper, a convex polygon composed of feature points is obtained by using the principle of convex hull. This polygon contains all the feature points on the edge of the object. Since the longest straight segment is used to reduce the measurement error, the side length of similar triangle is assumed to be the longest line segment. In this case, the straight line segment required for the measurement can be obtained.

## V. EXPERIMENT AND ANALYSIS

Figure 6 is the result of obtaining the image. Experimental images obtained by monocular camera were obtained at object distance d and u plus d respectively. Assuming the object distance was u, (a) images were obtained. Keep the parameters of the camera unchanged and move d equal to 100mm along the optical axis to obtain (b).The grabcut image segmentation method is used to segment the target object from the background, so that the target object can be emphasized and figure 7 is obtained. Figure 8 and figure 9 are the results of corner detection and feature point matching of the segmented image according to the improved Harris-SIFT algorithm. Figure 10 uses convex hull principle and similar triangle method to find out the line segment needed for calculation. Finally, the depth of the object is calculated by using the formula of object imaging principle.

(a) (b)

Figure 6. (a)Image obtained by object distanced (b)Iimage obtained at *u* plus *d*

Figure 7. Image segmentation

Figure 8. Corner detection

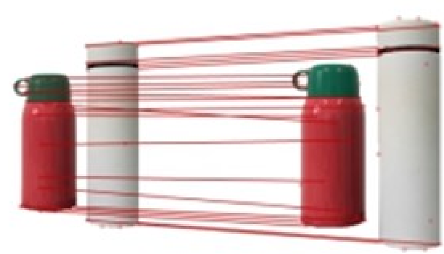


Figure 9. Feature point matching

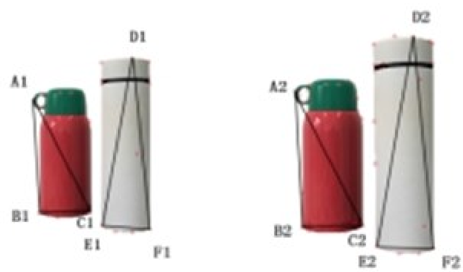



Figure 10. Select a straight segment

## VI. CONCLUSION

The object depth measurement method proposed in this paper can be acquired by ordinary camera. Easy to operate, no need to make any changes to the camera, do not need to calibrate the camera and other work. The depth measurement of the static question in the image can be completed automatically without manual intervention and verified by experiments. In this paper, based on feature point detection, the measurement accuracy of the depth information of the object in the image is high, which has a good use value.

A large number of experiments and comparative experiments have shown that the algorithm proposed in this chapter is simple and easy to operate with high measurement accuracy, but it also has the following shortcomings. As SIFT feature points need to be used in the measurement, this algorithm is not applicable to the objects with few SIFT feature points.

## REFERENCES

[1] Moravec H . Towards Automatic Visual Obstacle Avoidance[C]// International Joint Conference on Artificial Intelligence. Morgan Kaufmann Publishers Inc. 1977.

[2] Harris, C.G. and Stephens, M. (1988) A combined corner and edge detector. Proceedings of Fourth Alvey Vision Conference, 147-151.

[3] Lowe D G. Distinctive Image Features from Scale-Invariant Keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91-110.

[4] Lowe D G. Object recognition from local scale-invariant features[C]// Iccv. 1999.

[5] Liu Q , Yang Y , Gao Y , et al. A Bayesian framework for dense depth estimation based on spatial–temporal correlation[J]. Neurocomputing, 2013, 104:1-9.

[6] Sellent A , Favaro P . Optimized aperture shapes for depth estimation[M]. Elsevier Science Inc. 2014.

[7] Sun Z L , Lam K M , Gao Q W . Depth Estimation of Face Images Using the Nonlinear Least-Squares Model[J]. IEEE Transactions on Image Processing, 2012, 22(1):17-30.