IOP Conference Series: Earth and Environmental Science

PAPER • OPEN ACCESS

Ranging Method of Binocular Stereo Vision Based on Random Ferns and NCC Template Matching

To cite this article: Weiping Ma *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **252** 022149

IOP Conf. Series: Earth and Environmental Science **252** (2019) 022149 doi:10.1088/1755-1315/252/2/022149

Ranging Method of Binocular Stereo Vision Based on Random Ferns and NCC Template Matching

Weiping Ma[a], Wenxin Li[b] and Pengxia Cao[c]

Lanzhou Institute of Physics, China Academy of Space Technology, Lanzhou 730000, China.

*Corresponding author e-mail r: 498938802@qq.com [b]lwxcast@21cn.com [c]316657294@qq.com

Abstract. As an important research direction in the field of computer vision, binocular stereo vision has important application value for the reconstruction of scene 3D geometric information. A fast and accurate binocular vision target ranging method is proposed. Firstly, the random fern algorithm is used to accurately and quickly detect the left image target area to obtain the target center point. Then the NCC template matching algorithm is used to find the matching point in the right image. Finally, the two-dimensional coordinates of the matching point pairs of center point is used to calculate the distance between the target and the camera. The simulation results show that the ranging scheme can accurately measure the target distance within a certain distance range, and can achieve fast and effective target ranging.

1. Introduction

With the deepening of robot autonomy, the requirements of robot vision technology, which is the most important for robot autonomy, are also increasing. Binocular stereo vision is a key component of self-service robot technology. How to use the images captured by binocular camera to select the appropriate target Calibration points and acquire the target space location has always been the key and difficult point of binocular stereo vision technology. A common method is to rely on certain features of the target to detect the target, and use the centroid of the object as the calibration point. In this paper, the centroid is the target calibration point, and the random ferns feature matching algorithm is used for target detection. In a more complicated environment, the target area can be detected from the image to be tested, and then the NCC template matching method is used to realize the acquisition of the calibration points in the middle left and right images of binocular stereo vision system, and finally the position of the target calibration points is obtained by using the three-dimensional measurement principle.

2. Target detection based on random ferns

The random ferns algorithm is a machine learning classification method based on the random forest algorithm. It belongs to the semi-Naive Bayes classifier, which combines multiple features of the target into a random fern [1-2]. The random ferns algorithm has invariance to scale change, rotation, affine, illumination and occlusion, and avoids the huge amount of computation of feature invariance description and matching in traditional feature matching methods, which can improve the real-time performance of the system. The specific implementation process is shown in Figure 1. Shown.

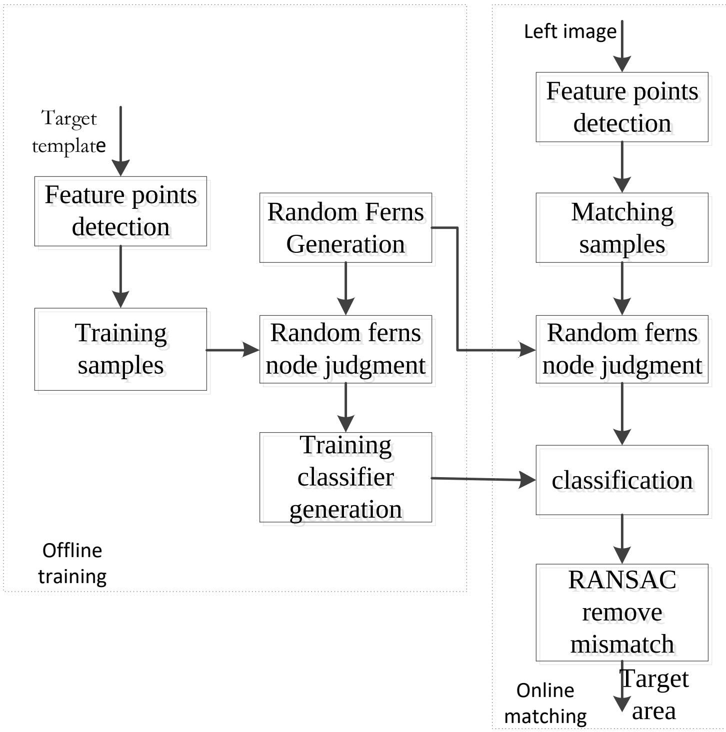

Figure 1. The process of Random Ferns Target detection

2.1. Offline training

The purpose of offline training is to generate a target classifier. Firstly, a certain number of feature points of the target are extracted, and some of the stable points are selected by affine transformation of the feature points, and then the stable feature points are taken as the center, and the binary features are extracted in the 32*32 neighbourhood as the feature attributes of the feature points. Finally, the posterior probability of each category feature attribute is calculated to construct the target classifier.

(1) Stable feature points extraction

In the random ferns algorithm, each feature point of the target template is regarded as a category, and the set of stable feature points constitutes the category set of the target. Therefore, it is important to extract stable feature points. The extraction steps are as follows: ①Extract the feature points set of the target template; ②perform multiple different affine transformations on the target template, extract the feature points in the image after each affine transformation; ③perform corresponding inverse affine transformation on these feature points to find the corresponding points in the feature points set of the target template; ④count the number of occurrences of each feature point in all affine transformations, and select the points with the most occurrences as the stable feature points.

(2) Training patch acquisition

To train a robust robust fern classifier, a large number of images of different scales and perspectives are required to form a training sample set. The random ferns also uses the affine transformation to generate all the training patches for each stable point. The specific implementation is to perform the affine transformation of the random parameters of the target template, centering on the corresponding points of the stable points in the affine image. 32*32 neighborhood pixels are used as training patches to synthesize sample sets of each category.

(3) Offline classifier generation

IOP Conf. Series: Earth and Environmental Science **252** (2019) 022149 doi:10.1088/1755-1315/252/2/022149

The random ferns algorithm treats all the training elements of the same stable point as the same class, and obtains the binary feature as the feature attribute set of the category by randomly selecting the test point pairs within the stable point patch [3]. Let the number of stable points be N, and the set of categories of the target be $C=\{c_i, i=1,2,\cdots,N\}$. The formula for calculating the characteristics is as follows:

$$f_j=\begin{cases}1 & if I(d_{j1})<I(d_{j2})\\ 0 & otherwise\end{cases} \tag{1}$$

Here, I represents the training patch (P) of a certain stable point, d_{j1} and d_{j2} represents arbitrary two pixel points (test point pairs) of the patch,the feature attribute set of the patch is $f=\{f_j, j=1,\cdots,M\}$. The purpose of the classifier is to assign classes for p,can be described by

$$\hat{c}_i=\arg\max_{c_i} P(C=c_i \mid f_1, f_2,\cdots, f_M) \tag{2}$$

According to the naive Bayesian classification model, equation (2) can be derived as follows:

$$P(C=c_i \mid f_1, f_2,\cdots, f_M)=\frac{P(f_1, f_2,\cdots, f_M \mid C=c_i)P(C=c_i)}{P(f_1, f_2,\cdots, f_M)} \tag{3}$$

The molecule is evenly distributed, and the denominator is a constant independent of the class, so the maximum of the formula (3) is expressed as

$$\hat{c}_i=\arg\max_{c_i} P(f_1, f_2,\cdots, f_M \mid C=c_i) \tag{4}$$

If it is assumed that the feature attributes are completely independent, the joint probability in the above formula can be expressed as the product of the conditional probability of each feature attribute to the class. However, considering the mutual relationship between image pixels, the complete independence between feature attributes does not exist, so the feature attribute set is divided. The binary features string M is divided into K groups, the length of every group is Forming a random ferns. The nodes in the fern are connected to each other, and the ferns and ferns are independent of each other. Equation (4) can be further simplified to

$$P(f_1, f_2,\cdots, f_M \mid C=c_i)=\prod_{k=1}^{K} P(F_k \mid C=c_i) \tag{5}$$

From the above training process for the random ferns classifier, the key to offline classifier training is to estimate the conditional probability of each fern for each category by calculating the frequency of

Occurrence of each category in the training sample $P(F_k \mid C=c_i)$, and the classifier is finally generated.

$$P(F_k \left| C = c_i \right.) = \frac{n_{k,i} + \sigma}{\sum_k (n_{k,i} + \sigma)} \tag{6}$$

Where, $n_{k,i}$ represents the number of samples of the category c_i, appearing in the *K*th fern, and the denominator is the number of all samples of the category c_i, σ is non-zero coefficient, generally taken as 1.

2.2. Online matching

In the online matching process, the feature points of the image to be matched are extracted, and 32*32 patches of the feature point are taken as the object to be classified, and the feature attribute set is obtained.The probability that the patch belongs to each category is counted by using equation (2), and the category with the largest posterior probability is the category to which the fragment belongs. After the feature points in the matched image are classified, the matching between the target template feature points and the image feature points to be matched is realized, and the target detection is completed.

3. NCC Template matching

Template matching [4] is a process of finding a known template image area in a search image, and is a matching method based on image gray scale. In the left and right images, the small range neighborhoods of the two corresponding points have similar gray distributions. Therefore, the correlation between the template image and the search image sub graph is used to determine the matching relationship between the corresponding points. The basic idea [5] is: centering on the target center point $O_l(x, y)$ determined in the left image $I_l(x, y)$, select the neighbourhood with the size $m \times n$ as the template *T*, and move the template *T* in the right image $I_r(x, y)$. The covered subgraph is *P*, the similarity measure function is used to discriminate the similarity of *T* and *P*. The maximum similarity is the best match, the center point $O_r(x, y)$ of the subgraph *P* is the corresponding point of $O_l(x, y)$.

Among the commonly used similarity measure functions, the normalized cross-correlation algorithm NCC has been proved to have good accuracy and adaptability. According to the polar line constraint condition in the binocular stereo matching, the ordinates of the corresponding points in the left and right images are equal, that is $y_l = y_r$, the search space is reduced to one dimension, and the search area is locked to all the pixels in the range y=y_l. The normalized cross-correlation algorithm is used to achieve the matching between the template center point and these pixels. The calculation formula of the NCC matching algorithm [6] is

$$R_{NCC} = \frac{\sum_m \sum_n (P(i,j) - \overline{P})(T(i,j) - \overline{T})}{\sqrt{\sum_m \sum_n (P(i,j) - \overline{P})^2} \sqrt{\sum_m \sum_n (T(i,j) - \overline{T})^2}} \tag{7}$$

where, $\overline{T}$ and $\overline{P}$ respectively represent the grayscale mean values of the left graph template T and the overlay subgraph P in the right graph, $T(i, j)$ and $P(i, j)$ respectively are a certain pixel point, and finally the largest value of R_{NCC} is the best matching block whose center point is the matching center point.Then the calibration points of the left and right images are successfully obtained.

4. Binocular visual calibration point acquisition

Binocular stereo vision [7] is a simulation of human eyes. The eyes observe three-dimensional objects from two different angles. According to the projection principle of geometric optics, the image points are located at different positions on the left and right retinas. The difference in position between the retinas of both eyes is the binocular parallax, which reflects the depth information of the objective object. The binocular stereo vision uses the parallax theory to image the same object at different positions with two cameras, acquire the parallax of a certain point, and finally calculate the depth information of the object by using the parallax ranging principle. The most commonly used in the binocular vision system is the parallel binocular stereo vision system. The two cameras of the system are placed in parallel, and the specific model [8] is shown in Figure 2.

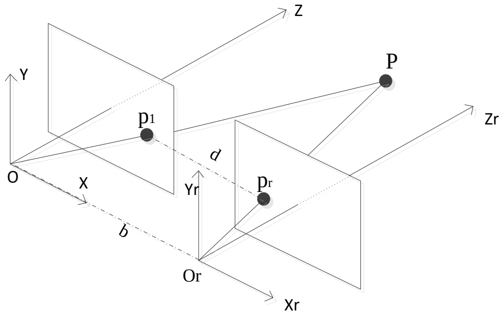

Figure 2. Parallel binocular stereo vision model

Let the spatial point P, its projection points of the left and right imaging plane coordinate systems are $p_l(x_l, y_l)$ and $p_r(x_r, y_r)$ respectively, where $y = y_l = y_r$, the parallax is defined as $d = x_l - x_r$, according to Figure 2 there are:

$$\frac{b-(x_l - x_r)}{Z-f} = \frac{b}{Z} \Rightarrow Z = \frac{bf}{d} \tag{8}$$

It can be seen from the above formula that as long as the matching point pair coordinates, the camera focal length and the baseline are known, the corresponding spatial point three-dimensional coordinates can be obtained.

5. Binocular ranging test

In order to verify the feasibility of the algorithm, the target ranging test was carried out based on the constructed binocular vision system. Firstly, the strong applicability of the random ferns algorithm is verified. The target is detected in the scene with occlusion, rotation and scaling. The result is shown in Figure 3. Secondly, the target is placed at different distances from the binocular camera. The object is determined by the distance measurement algorithm [9] of this paper, and its accuracy is analyzed by comparing with the real measurement distance. The measured distance is the vertical distance from the closest point of the target object to the center point of the camera baseline.

(a) (b) (c) (d)

Figure 3. The target (a) the random ferns target detection on occlusion (b), rotation(c) and scaling (d)

Set the measured distance to 4 different values, such as 100~300mm, at intervals of 50mm. The measured data are shown in Table 1.

Table 1. The Comparison of actual and calculated Distance from cameres to target

Actual Distance (mm)	Left Center Point	Right Center point	Calculated Distance (mm)	Error (mm)
100	(210,-49)	(-182,-49)	107	7
150	(220,-38)	(-44,-38)	159	9
200	(107,20)	(-108,20)	195	5
250	(143,-91)	(-30,-91)	242	8

It can be seen that the ranging algorithm in this paper can accurately measure the target distance, meet the accuracy requirements of the ranging device, and maintain high ranging accuracy with1n a certain range. However, there are still measurement errors, and there are three main reasons for the analysis error:

1) The binocular camera fails to achieve a completely ideal parallel binocular model, and there is a placement error between the left and right cameras, which may result in inaccurate results;

2) The image collection quality is not high, which affects the measurement results;

3) The calibration error of the internal and external parameters of the camera may cause deviations in the acquisition of depth information.

6. Conclusion

A binocular vision ranging algorithm is proposed. The target detection is realized by random fern feature matching algorithm. Then the NCC template matching algorithm is used to obtain the calibration points of the left and right images. Finally, the distance between the target and the camera is calculated by using the three-dimensional measurement principle. Binocular vision ranging. The positioning method takes into account the accuracy and real-time performance, and can meet the distance measurement requirements within a certain range.

References

[1] Ozuysal M,Calonder M, Lepetit V, et al. FAST keypoint recognition using random ferns [J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2010,32(3):448-461.
[2] Luo Y, Fu Y, Zhang Y. A Monocular-vision Real-time Matching Algorithm Based on Fast Corners and Affine-improved Random Ferns [J]. ROBOT, 2014, 36(3):271-278.
[3] Sun B W, Qiu Z J, Shen B, et al. Random Ferns Feature Matching Algorithm Based on Orientation Information [J].Computer Engineering,2014,40(5):192-195,202.
[4] Shinichiro O, Masako O. Fast template matching with polynomials [J]. IEEE Transactions on Image Processing,2007, 16(8):2139-2149.
[5] Jia D, Yang N H, Sun J G. Template selection and matching algorithm for image matching [J].Journal of Image and Graphics,2017,22(11) :1512-1520.
[6] Wu P ,Wang B, Xue J, et al. Research of Accelerate Detect of Pulmonary Nodules Based on Template Algorithm [J].Computer Engineering and Applications,2015,51(7):184-188.
[7] Cao Z L, Yan Z H, Wang H. Summary of Binocular Stereo Vision Matching Technology [J].Journal of Chongqing University of Technology(Natural Science),2015,29(2):70-75.
[8] Wang B, Guo L. 3D Measurement Technology Based on Binocular Stereo Vision Using Gary Code and Multi-Step Phase-Shift Structured Light [J].Computer Measurement& Control,2018,26(5):25-28.
[9] Li Q M, Li Y Y. Binocular Stereo Distance Measurement Based on Improved SIFT Algorithm [J].Transducer and Microsystem Technologies,2017,36(11):119-122.