Research on obstacle detection and path planning based on visual navigation for mobile robot

Yixiao Zheng¹, Bixi Yan¹, Chao Ma², Xinyu Wang¹, Haoyang Xue¹

¹ School of Instrument Science and Optoelectronic Engineering, Beijing Information Science and Technology University, Beijing, China
² Beijing Key Laboratory of Intelligent Space Robotic Systems Technology and Applications, Beijing Institute of Spacecraft System Engineering, Beijing, China
beixinshipei@163.com

Abstract. Both obstacle and path planning are of significance for autonomous visual navigation of a mobile robot which is able to realize free movement and obstacle avoidance. In this paper, a scheme of obstacle detection which is based on binocular vision to identify obstacles is constructed and it is combined with improved A* algorithm for obstacle avoidance motion for mobile robot platform. First, a binocular vision system to detect the robot's perception of the environment is used and the visual information is obtained. Then a Gaussian filter is used to reduce the noise in the image, and the grayscale image is obtained through the binocular vision system. The gray scale map is matched with SAD (Sum of Absolute Differences) in order to obtain the disparity image with comparison. Second, the location information of the obstacle is extracted through the environmental depth information contained in the disparity image, and the improved A* algorithm is used as the path planning algorithm in this paper, which can modify the path of the planning error. Furthermore, the detection experiment of the robot visual obstacle avoidance system was established and the optimal path was obtained. The examples show that, the obstacle avoidance system can complete the task of obstacle detection and give a better path for the mobile robot.

1. Introduction

With the continuous development of the detection technology, the future development direction of robot in our country is required by the following deep exploration task. The ability of robot's environment perception becomes more and more important. A necessary part of environmental perception is the detection of obstacles. Because of the unknown detection environment, higher requirements are put forward for the obstacle avoidance ability of the robot.

The obstacle detection system is the premise for the robot to judge the external environment and plan the path, it is also the premise that the robot can successfully complete the task. At present, there are four main methods of obstacle detection, LIDAR detection, machine vision detection, infrared sensor detection, ultrasonic sensor detection. Both ultrasonic and infrared sensors are only suitable for close range detection, and only for the distance measurement of a single obstacle [¹]. The appearance of laser sensors solves the problem that these two kinds of sensors have a small range of ranging and can accurately obtain the position information of objects. But it has not been fully applied to laser range finders because of the high price and large size[²]. Compared with other sensors, visual sensors have the advantages of high test accuracy, strong autonomy, high reliability and large information
capacity\textsuperscript{[3]}. Meanwhile it has the characteristics of small size, light weight, low power consumption, long life and high cost performance\textsuperscript{[4]}.

The path planning of mobile robot is based on the path planning of sensor information, which is unknown to the size, shape and position of obstacles, and can only be obtained by sensors. The path planning problem has been deeply studied by many scholars at home and abroad for more than 20 years, and has produced a lot of research results. Several path planning methods are described below\textsuperscript{[5]}.

The visual map method connects the obstacle points in the map with the target points of the robot, removes the line through the obstacle, and the remaining line is the visual map we need. The principle of artificial potential field method is that the robot can avoid obstacles and plan its path by the repulsive force around obstacles and the attraction of target points. A* algorithm is the most effective direct search method to solve the shortest path in static road network, and it is also a common heuristic algorithm for many other problems. A* algorithm must find the shortest path and high efficiency.\textsuperscript{[6]}

2. Framework

This system mainly includes two functional modules, namely obstacle detection and path planning. First, the images on the left and right sides of the binocular camera are collected, the camera is calibrated, and then the obstacles in the image are detected by using the Bumblebee 2 through software processing, and the three-dimensional information of the obstacle is obtained through the disparity image. Input the information into the path planning module, so as to avoid obstacles and plan the forward path.

![Image](image1.png)

Figure 1. System framework diagram

3. Obstacle detection

In this paper, we use the method of obstacle detection in unknown environment by disparity image, and detect obstacles by region connectivity and threshold filtering through disparity image \textsuperscript{[7]}.

3.1. Stereo Vision

The standard stereo vision system is composed of two cameras with strictly parallel optical axes, as shown in figure 2. The image planes of the two cameras are precisely located in the same plane, with a certain base distance $T$ and the same focal length $f$.

![Image](image2.png)

Figure 2. Stereo vision system

Suppose any point in the physical world $P$ the imaging point on the left and right images are $P_1$ and $P_2$. The corresponding transverse coordinates and vertical coordinates are $x_1$ and $x_2$, $d=x_1-x_2$ is defined as parallax, respectively, then the distance from the $P$ to the center line of the camera is $Z$\textsuperscript{[8]}.

$$Z = \frac{fT}{x_1-x_2} = \frac{fT}{d}$$  \hspace{1cm} (1)
Camera calibration is the primary task of stereo vision. In this paper, the camera calibration method proposed by professor Zhang in 1998 is used to obtain the internal and external parameters of the camera. After using the internal parameters of the camera, the camera is distorted and the polar line is calibrated, and the matching points can be found from different viewpoint images to complete the stereo matching.

Monocular calibration is the premise of camera binocular calibration. Only after the internal and external parameters of the two cameras are known, the binocular calibration can be carried out. First, the coordinate projection model is established to clarify how points in reality are mapped to the image plane through rotation and translation in the world coordinate system, as shown in Figure 3.

**Figure 3. Coordinate relation transformation**

The transformation from the world coordinate system \((X_w, Y_w, Z_w)\) to the camera coordinate system \((X_c, Y_c, Z_c)\) can be expressed as:

\[
\begin{bmatrix}
X_c \\
Y_c \\
Z_c
\end{bmatrix} =
\begin{bmatrix}
R & T \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
X_w \\
Y_w \\
Z_w
\end{bmatrix}
\]

in the formula: \(R\)—3×3 Orthogonal transformation matrix; 
\(T\)—Three-dimensional translation vector, \(0=\langle 0 \ 0 \ 0 \rangle\).

The transformation from the camera coordinate system \((X_c, Y_c, Z_c)\) to the image plane coordinate system \((x, y)\), according to the pinhole imaging model, the geometric relationship can be obtained:

\[
f = \frac{x}{Z_c} = \frac{y}{X_c}
\]

\(f\) is the focal length of the camera and is represented by a matrix of homogeneous coordinates:

\[
\begin{bmatrix}
x \\
y \\
1
\end{bmatrix} =
\begin{bmatrix}
f & 0 & 0 & 0 \\
0 & f & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
X_c \\
Y_c \\
Z_c
\end{bmatrix}
\]

The transformation from the image plane coordinates \((x, y)\) to the image pixel coordinates \((u, v)\) is homogeneous coordinate matrix:

\[
\begin{bmatrix}
u \\
v \\
1
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & u_0 \\
0 & 1 & v_0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
1
\end{bmatrix}
\]

in the formula: \(dx, dy\) — the physical dimensions of pixels in the \(x\) and \(y\) axis directions. \((u_0, v_0)\) is the origin of the image plane coordinate system. The following three forms are obtained:
\[
Z_c \begin{bmatrix}
u \\
v \\
1
\end{bmatrix} = \begin{bmatrix}
\frac{1}{dx} & 0 & u_0 \\
0 & \frac{1}{dy} & v_0 \\
0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
f & 0 & 0 & 0 \\
0 & f & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix} \begin{bmatrix}
X_w \\
Y_w \\
Z_w \\
1
\end{bmatrix} = M_1 M_2 \begin{bmatrix}
X_w \\
Y_w \\
Z_w \\
1
\end{bmatrix} = M \begin{bmatrix}
X_w \\
Y_w \\
Z_w \\
1
\end{bmatrix}
\] (5)

in the formula:

\[
a_s = \frac{f}{dx}, a_v = \frac{f}{dy}, M_1 = \begin{bmatrix}
a_s & 0 & u_0 & 0 \\
0 & a_v & v_0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}, M_2 = \begin{bmatrix}
R & T \\
0^T & 1
\end{bmatrix}
\]

\(a_s\) and \(a_v\) are the normalized focal lengths of the \(u\) and \(v\) axes, respectively, and \(M = M_1 M_2\) is the projection matrix of the camera. \(M_1\) is the camera internal parameter, \(M_2\) is the camera external parameter. The process of camera calibration is essentially the process of determining these internal and external parameters.

When the left and right camera is calibrated, the relative position between the left and right cameras needs to be known, so two matrices are introduced, which are the rotation matrix \(S\), and the translation matrix \(T\). The relationship between the two cameras is:

\[
\begin{bmatrix}
X_{cl} \\
Y_{cl} \\
Z_{cl}
\end{bmatrix} = S \begin{bmatrix}
X_{cr} \\
Y_{cr} \\
Z_{cr}
\end{bmatrix} + T
\] (6)

After the internal and external parameters of the left and right cameras are obtained according to the monocular calibration method, it can be concluded that:

\[
Z_{cl} \begin{bmatrix} u_l \\
v_l \\
1
\end{bmatrix} = M_1 M_2 \begin{bmatrix}
X_w \\
Y_w \\
Z_w \\
1
\end{bmatrix}, Z_{cr} \begin{bmatrix} u_r \\
v_r \\
1
\end{bmatrix} = M_1 M_2 \begin{bmatrix}
X_w \\
Y_w \\
Z_w \\
1
\end{bmatrix}
\] (7)

According to the checkerboard calibration checkerboard, the rotation matrix \(R\) and the translation matrix \(T\) can be obtained by existing applications.

After obtaining the inside and outside parameters of the camera, the binocular image is then corrected. There are two correction steps. One is to eliminate the distortion of the lens to the image, which can be eliminated according to the camera model after obtaining the camera parameters by the camera calibration above; the other is to make the binocular image coplanar and strictly aligned, and the binocular image can be re-projected by Bouget re-projection algorithm.

### 3.2. Obstacle Detection

Because there are some reasons to cause some noise in the image, and these noises will seriously affect the next processing. We will smooth the image to reduce the noise. Gauss filtering method is a linear smoothing filtering method, which is generally used in the denoising process of image processing. Gauss filtering is the process of weighted average of the whole image, and the value of each pixel is obtained by weighted average of its own and other pixel values in the neighbourhood. The Gauss filtering method is fast and the effect is ideal.

Then match the processed image. Binocular matching is to match a projection point on an object on two images taken at the same time in the left and right cameras, which is to match the target point on one image to the target point on the other image. Based on the polar constraints, the energy function
reflecting the matching relationship between pixels is constructed. Then, by solving the extreme value of the energy function, the matching relationship between each pixel point on the binocular camera image is obtained, so as to realize the stereo matching of the binocular vision system. The output is the disparity image of the binocular image.

Because the disparity image reflects the depth information in the scene, the target with the same depth corresponds to the same gray value on the disparity image. Therefore, for the forward flat road, the disparity image should present an approximate corrugated shape in the horizontal direction, and as the depth increases, the image brightness gradually decreases, and when there is an obstacle at a certain distance, the corrugated shape of the position changes significantly. By scanning the disparity image one by one, we can determine the position where the disparity image changes, that is, the position of the obstacle on the disparity image, so as to realize the detection of the obstacle, and then carry on the 3D reconstruction to the extracted area, then we can determine the obstacle.

4. Path planning

When we are given the starting point and the target point position information, a best route is planned between two points. Common path prediction algorithms include breath-first search (BFS) algorithm, Dijkstra algorithm and A* search algorithm. The shortest path algorithm aims to find the shortest path between nodes, which not only ensures the optimal, but also reduces the time complexity and space complexity.

4.1. Principle

A* algorithm introduces a valuation function based on breadth-first search, which is used to evaluate all expandable points every time, so as to find the optimal expandable point, and then search from that point until the target point is found. The valuation function is expressed as follows\(^{(12)}\):

\[
f_n = g_n + h_n \tag{8}
\]

\(f_n\) —evaluation function n from initial point through intermediate node to target point

\(g_n\) —actual cost of initial point n to target point through intermediate node

\(h_n\) —estimated cost of n the best path to the target point

The selection of estimation cost \(h_n\) is very important, \(h_n\) the closer to the actual distance, the better the evaluation function results. The distance is estimated to be two points between Manhattan distance, Euclidean distance, Chebyshev distance and so on.

Manhattan distance:

\[
D = |(x_1 - x_2)| + |(y_1 - y_2)| \tag{9}
\]

Euclidean distance:

\[
D = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{10}
\]

Chebyshev distance

\[
D = \max \left(|x_1 - x_2|, |y_1 - y_2|\right) \tag{11}
\]

4.2. Work Flow

When \(g_n\) and \(h_n\) are known, A* algorithm takes the following steps\(^{(13)}\):

Add the start node to the open list; Traverses the open list to find the node that minimizes the valuation function, which is the node that is currently being processed. Put this node in the closed list, for this node adjacent to the top, bottom, left, right, upper left, lower left, upper right, lower right and so on. If they are not reachable or exist in the closed list, ignore it, otherwise, do the following: If it is not in the open list, add it to the open list and set the current node to its parent, and record the \(f_n, g_n\) and \(h_n\) values of the node. If it is already in the open list, check that there is a current node to it. If \(g_n\) value is smaller, it means this is a better path. If so, set its parent to the current node and recalculate its\( g_n\).
and $h_n$ values. When all the target points are added to the open list, the shortest path is found. Starting from the target point, moving along the parent node to the starting point is the resulting path.

4.3. Optimization

The approach used in this article is to recalculate the $g$ and $f$ values regardless of whether the expansion point has entered the open list, and then select the node with the lowest $f$ value. The embodiment of path planning is that every step is optimal, and this step is based on the previous step, and the closed list is returned through step by step optimization.

Valuation function:

\[
f_n = g_n + h_n
\]

\[
g_n = \begin{cases} 
g_{n-1} + \alpha, \\
g_{n-1} + \beta,
\end{cases}
\]

$g_{n-1} + \alpha$ is adjacent nodes along the diagonal, $g_{n-1} + \beta$ is adjacent nodes in the horizontal and vertical direction, $\alpha=1.4$, $\beta=1.0$.

5. Experiments

5.1. Camera Calibration

In this paper, Bumblebee 2 camera is selected as the tool for image acquisition. The depth range of this camera is 0.3-25 m, and the resolution is 2208*1242. Figure 4 shows the robot Navigator Q2 with camera.

![Figure 4. The robot Navigator Q2](image)

According to the 10 images collected by the camera, the range of corner points is determined, and the intersection points are extracted, and the parameters are calculated at the same time.

![Figure 5. Camera calibration grid](image)

Select the yellow point in the upper left corner as the origin, draw the camera coordinate system, automatically calculate the number of squares and corner points ($X$ direction and $Y$ direction), and show that the corner points will also be extracted.
5.2. Picture Processing
Gaussian filter is used to smooth the left and right camera images, which can reduce the noise in the images and obtain clearer images. At the same time, gray processing and correction of the image, gray map to obtain convenient SAD matching.

| Image                        | Left camera | Right camera |
|------------------------------|-------------|--------------|
| Original image               | ![Original Image](image1.png) | ![Original Image](image2.png) |
| Gaussian filter for gray-scale image | ![Gaussian Filter Image](image3.png) | ![Gaussian Filter Image](image4.png) |
| Corrected image             | ![Corrected Image](image5.png) | ![Corrected Image](image6.png) |

5.3. Disparity image Acquisition
To get an accurate disparity image, you need to set an appropriate SAD window, which will neither select the background box of the scene as an obstacle nor select the wrong obstacle. Below compare the different sizes of SAD window, select an appropriate value for the next step of recognition.
Table 2. Comparison of disparity images of different window sizes

| SAD window size (unit window size: 16) | Disparity image | SAD window size (unit window size: 16) | Disparity image |
|---------------------------------------|-----------------|---------------------------------------|-----------------|
| 1*16                                  |                 | 2*16                                  |                 |
| 3*16                                  |                 | 4*16                                  |                 |
| 5*16                                  |                 | 6*16                                  |                 |
| 7*16                                  |                 | 8*16                                  |                 |

According to the comparison, a more appropriate disparity image can be obtained when the window size is 3*16. The disparity image can provide more environmental information without confusing the background with the obstacle, so as to obtain more accurate location of the obstacle.

The processed images are put into the software to obtain the disparity image, we can see that the other scenes in the obstacle vision map are filtered out, and the obstacles are extracted.

5.4. Obstacle Coordinate Acquisition
The mouse is used to click on the obstacle in the disparity image, the coordinates of the obstacle points are displayed in the window, and the image coordinates are transformed into world coordinates.
Table 3. Obstacle coordinates

| Image coordinates | World coordinates |
|-------------------|-------------------|
| (1242,161)        | (-219.213,594.88,-1784.87) |
| (1250,241)        | (-217.443,438.184,-1522.06) |
| (1004,163)        | (86.1296,626.576,-1887.96) |
| (1090,152)        | (-44.1173,610.675,-1729.16) |
| (185,487)         | (1322.92,222.76,-2072.8) |
| (190,499)         | (-105418, -21965.5,160000) |
| (540,485)         | (1007.37,290.413,-2665.93) |
| (619,605)         | (723.375,42.3594,1796.69) |

5.5. Path Planning

According to the coordinates, a map is built, determining the starting point and the end point, improved A* algorithm trend robot to find the shortest path that can reach the end point.

Figure 8. An improved A* algorithm path planning diagram with opposite starting point

According to the comparison of the figure above, it can be found that when the obstacles are the same, different starting and ending points will also affect the trajectory of path planning.

The coordinates of the obstacles in the table can be obtained by the obstacle detection system and set on the map. Then, the improved A* algorithm is used to calculate the estimated cost of each point and form a path after comparison.

6. Conclusion

In this paper, the obstacle detection based on stereo vision and the path planning based on improved A* algorithm was researched. In order to complete obstacle detection, the principle of stereo vision system, camera calibration, and disparity image acquisition technologies are studied for mobile robot platform. Compared with other obstacle detection, Gaussian filter is used to process and smooth the image, which is more conducive to obtaining the disparity image. Some functions were used to process the image to obtain the disparity image and obtain the location information of the obstacle. The principle and flow of the improved A* algorithm was studied and used to realize the path planning. In this paper, the robot obstacle avoidance system combined with stereo vision obstacle detection and improved A* algorithm path planning can accurately identify the coordinates of
obstacles, establish a map, and plan the path of the detection robot. The examples show that, the obstacle avoidance system designed in this paper can complete the task of obstacle detection and give a better path for the mobile robot.

Acknowledgments
This research was supported, in part, by the National Natural Science Foundation of China (No. 51975044), the Science and Technology on Space Intelligent Control Laboratory (No. HTKJ2019KL502016) and 2019 Practice Training Program of Beijing Information Science and Technology University.

References

[1] Hou LC, Sun ZQ and Lu R 2007(02):7 Application of infrared sensor in robot obstacle avoidance system [J]. Science and Technology Advisory Guide.
[2] Cai ZX, Xiao Z, Yu JX 2008(02) Real-time detection of dynamic obstacles based on lidar [J]. Control Engineering 200-203.
[3] Tan QM, Wei QQ, Ma RQ, Chen L, Lin Y 2017: Study on visual navigation system for intelligent robot in the unstructured environment. 2017 IEEE International Conference on Robotics and Biomimetics.
[4] Li LY, Study on visual navigation system of forklift based on feature tracking [D], Zhengzhou University.
[5] Fang YQ 2009, (11) Current situation and development of mobile robot path planning [J]. Science and Technology Communication 67-68.
[6] Li LX, Li WG, Li ZX 2013(11) Fan Literature. Discussion and application of robot path planning method [J]. Mechanical Engr 8-9.
[7] Gou Q, Geng N, Zhang ZY 2013,34(2) Detection method of farmland obstacle in unknown environment based on disparity image [J]. Computer Engineering and Design 707-710.
[8] Lu XD, Wang MQ, Hao LH, Ding J 2017,31(06) A study on the measurement method of signal bomb height based on binocular stereo vision [J]. Journal of Testing Technology 505-511.
[9] Fang BW, Zhang XD, Chen JY and Sun JY 2019(04) A study on detection method of obstacle distance in driving based on binocular vision [J]. Mechanical Design and Manufacturing 94-98.
[10] Ruan QQ 2001 Digital Image Processing [M]. Electronic Industry Press.
[11] Ning KL 2014 Research and implementation of 3D reconstruction platform based on binocular stereo vision [D]. Beijing University of Posts and Telecommunications.
[12] Cheng J, Chen YJ and Liu ZF 2019,19(29) Path planning of unmanned surface craft based on A* algorithm [J]. Science, Technology and Engineering 201-206.
[13] Liu FF 2016,37(12) A* algorithm based on user feedback [J]. Neijiang Technology,41-42.