

# Automatic Target Positioning of Bomb-disposal Robot Based on Binocular Stereo Vision

Hengnian Qi<sup>1,2</sup> Wei Wang<sup>3</sup> Liangzhong Jiang<sup>1</sup> Luqiao Fan<sup>1</sup>

<sup>1</sup>School of Mechanical Engineering, South China University of Technology, Guangzhou 510641, P. R. China

<sup>2</sup>School of Information Engineering, Zhejiang Forestry University, Lin'an 311300, P. R. China

<sup>3</sup>College of Automation, South China University of Technology, Guangzhou 510641, P. R. China

## Abstract

Automatic target positioning is necessary for a bomb-disposal robot to perform the duty of disposing a perilous object. We proposed the design of a bomb-disposal robot based on binocular stereo vision which can achieve the state-of-the-art of automatic target positioning with high accuracy. In the intelligent robot system, the three dimensional coordinate of the perilous target is acquired using the binocular stereo vision technique, then the manipulator will be planned and controlled to reach the target according to the acquired three dimensional coordinate. To implement the design, the whole robot system can be decomposed into binocular stereo vision subsystem, and embedded motion control subsystem. The experiment result shows a satisfying level of accuracy in performing the task of automatic target positioning.

**Keywords:** Binocular stereo vision, Motion planning, Manipulator, Target positioning

## 1. Introduction

Several kinds of robots have been developed in different application areas in recent years [1], which benefit from the much research work that has been done in the academic communities of robotics and also the great requirement for such robots in various application fields. For example, the bomb-disposal robot, which is used to perform bomb disposal duties in an environment that is distant, hazardous or otherwise accepted as being inaccessible to humans, has been used widely by the police all over the world and has improved the safety for the duty of bomb disposal greatly.

Different types of bomb-disposal robots have been developed and used in different countries all over the world, for example the CYCLOPS which is produced by ABP company and MV4 produced by Telerob in Germany. These robots have extended the ability of human beings greatly to deal with the dangerous tasks of finding and defusing bombs. The

key point for a bomb-disposal robot to perform the duty of bomb disposal is target positioning, which implies to control the manipulator fixed on the robot to reach the position of the perilous target. However, the manipulator of current bomb-disposal robots are all designed to be manually controlled by an operator remotely, which bring in some disadvantages: the operator need to be well trained, high accuracy of target positioning cannot be achieved. So automatic operations or intelligent techniques are necessary for such robots.

Two key problems need to be investigated in order to reach the state-of-the-art of automatic target positioning. First, the three dimensional coordinate of the perilous target need to be acquired with high accuracy in the robot-based coordinate system, this issue can be settled through the binocular stereo vision technique. Another problem is the motion planning of the manipulator, which is to guide the manipulator to reach the target position automatically according to the accurate 3D coordinate acquired, instead of the manual controlling of the joints by the operator.

In this paper, the design of the whole bomb-disposal robot system is described in detail. The remainder of the paper is organized as follows. In section 2, we propose the architecture of the robot, including the mechanical configuration of the robot and the whole intelligent robot system configuration. In section 3, the design of the vision subsystem is described. In section 4, the embedded motion control subsystem is described, which takes charge of motion planning and low-level control of motors. Section 5 concludes the paper.

## 2. System architecture

### 2.1. Mechanical configuration

The structure of the bomb-disposal robot consists of a remotely operated vehicle and a manipulator fixed on it with five DOF(Degree of Freedom). This kind of

structure has been adopted in several commercial bomb-disposal robots and is considered as an effectual structure.

The vehicle will be remotely driven to get close to the target in an environment which is too distant or hazardous for a human expert. When the target is within the operating space of the manipulator, the manipulator will start up to perform its duties.

The manipulator which is built on the remote vehicle is the key component of the robot. The mechanical structure of the manipulator is designed according to the working principle of the arm of human beings. The five DOF of the manipulator, which come from the rotation joint of waist, the joint of shoulder, the joint of elbow, the joint of wrist, the rotate joint of claw, can provide a satisfying agility and manageability for the operation in various surroundings. The configuration of the manipulator is shown Fig. 1.

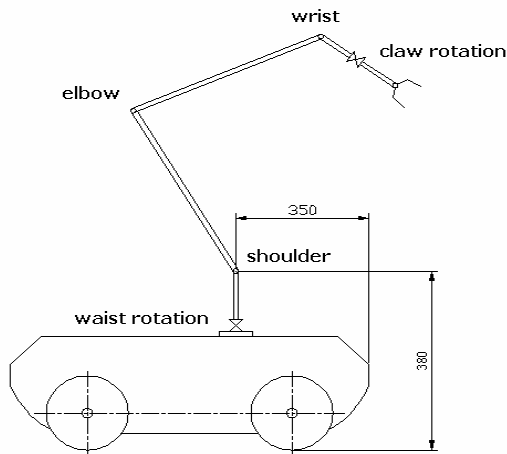


Fig. 1: Configuration of the manipulator with 5 DOF.

## 2.2. The robot system configuration

To make the robot implement automatic target positioning with binocular stereo vision technique, a remote server is needed. The remote server is actually a remote control center, which is considered to be the brain of the robot. It performs tasks such as monitoring the status of the robot, receiving the video acquired by cameras, sending control instructions to the robot, and also providing the man-machine interface. The intelligent level processing is also located on the server. The system architecture is shown in Fig. 2.

The whole robot system can be decomposed into binocular stereo vision subsystem, and embedded motion control subsystem. The binocular stereo vision subsystem is used to generate the 3d coordinate of the

target based on the images acquired by the cameras, while the embedded motion control subsystem is in charge of the locally controlling of the manipulator. The processing procedure of the whole robot system to perform the duty of disposing a target with automatic target positioning is as follows:

- 1> The vehicle is driven to the extent where the target is within the work space of the manipulator.
- 2> Start the vision system, the images of the target is sent to server.
- 3> The target is marked in the image by the operator through the human-user interface, and 3D coordinate of the target is generated through the two image coordinate of the target.
- 4> The camera-based 3d coordinate of the target is transformed into the robot-based 3d coordinate
- 5> Control instructions is generated based on the robot-based 3d coordinate of the target and then is sent to the embedded motion control subsystem
- 6> According to the control instructions received, the embedded motion control subsystem controls the manipulator to reach the position of the target.

The binocular stereo vision subsystem, and embedded motion control subsystem will be illustrated in details in the following sections.

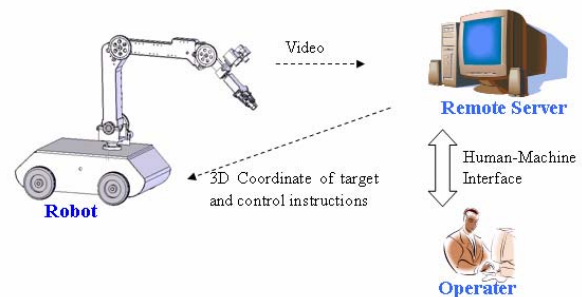


Fig. 2: System architecture of the whole robot system.

## 3. Binocular stereo vision subsystem

Binocular stereo vision is a well studied problem due to the much research work that has been done in the computer vision community in the past few years [2]-[5], especially the research work of Tasi [6] and Zhang [7] on the camera calibration technique, which is a necessary step in 3d computer vision in order to extract metric information from 2d images. The objective of our binocular stereo vision subsystem is to acquire the camera-based 3d coordinate of the target

utilizing 2d images. The images are captured from two cameras built on the manipulator and are sent to the remote server through a radio. In order to achieve the objective, two problems needed to be investigated: First, the projection mechanism of the camera from the real world to the 2d images need to be modeled and identified, this procedure is called calibration, including single camera calibration that calibrate a single camera and stereo calibration which is used to identify the relative location of the right camera with respect to the left one; another problem is called stereo triangulation which is used to generate the camera-based 3d coordinate utilizing the two 2d image coordinates of the target in the and camera and right camera and the result of the camera calibration procedure.

### 3.1. Camera calibration

In the procedure of single camera calibration, we adopted the calibration engine with Zhang's planar calibration technique[7] which can easily calibrate a camera. The technique only requires the camera to observe a planar pattern shown at a few (at least two) different orientations, it is very easy to calibrate, do not need the expensive calibration apparatus, and elaborate setup, which made it more flexible and more suitable for the usage in our application, the calibration images are shown in Fig.3. The proposed procedure consists of a closed-form solution, followed by a nonlinear refinement based on the maximum likelihood criterion, where very good result have been obtained through the experiment result. The calibration technique has been considered to advance 3D computer vision one step from laboratory environments to real world use.

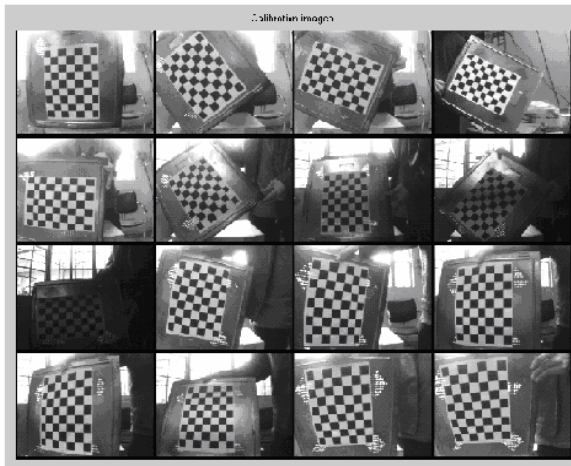


Fig. 3: Calibration images.

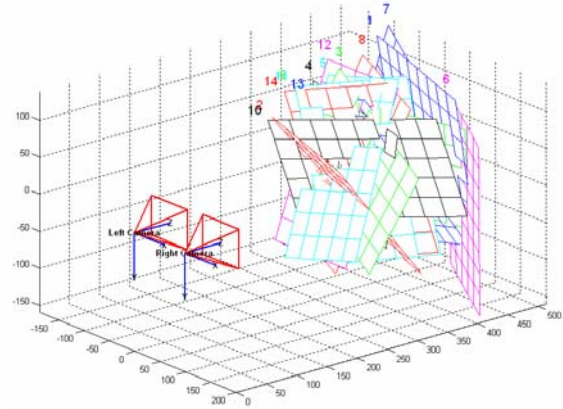


Fig. 4: Calibration result.

After the single camera calibration procedure, the projection model of the left and the right camera can be identified respectively as follows:

$$C_{left} = H_{left}(I_{left})$$

$$C_{right} = H_{right}(I_{right})$$

Where  $C_{left}$  is the camera-based 3d coordinate in the left camera coordinate system, which is centered at the optical center of the left camera, with the x axis the same as the optical axis.  $I_{left}$  is the image coordinate of the target in the left camera.  $H_{left}$  actually represents the projection model which is identified in the single camera calibration procedure.  $C_{right}$ ,  $I_{right}$  and  $H_{right}$  represent correspondent ones for the right camera.

After the two cameras are calibrated separately, the relative location of the right camera with respect to the left one can be identified in the stereo calibration procedure:

$$\begin{bmatrix} C_{right} \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \times \begin{bmatrix} C_{left} \\ 1 \end{bmatrix}$$

Where  $R$  is the  $3 \times 3$  rotation matrix and  $T$  is the translation vector which represent the transformation from the left camera coordinate system to the right camera coordinate system. In order to identify  $R$  and  $T$  in the stereo calibration procedure, a least-squares estimator is constructed to get the optimal estimation through a series of data of  $C_{left}$  and  $C_{right}$  acquired in the single camera calibration. The calibration result is shown in Fig.4.

### 3.2. Stereo triangulation

After the whole calibration procedure, the parameter of our binocular stereo vision model is identified and can be utilized to generate the camera-based 3d coordinate of a target through the following stereo triangulation algorithm.

$$\begin{cases} C_{left} = H_{left}(I_{left}) \\ R \times C_{left} + T = H_{right}(I_{right}) \end{cases}$$

Given  $I_{left}$  and  $I_{right}$ , the camera-based 3d coordinate of the target in the left camera coordinate system can be obtained by solving the equations above.

However, to identify a target and get the corresponding image coordinate  $I_{left}$  and  $I_{right}$  is a difficult problem. We can refer to human beings who can mark the target through the human-machine interface, see Fig.5, where a bottom is marked and the two image coordinates are acquired to generate the 3d camera based coordinate.

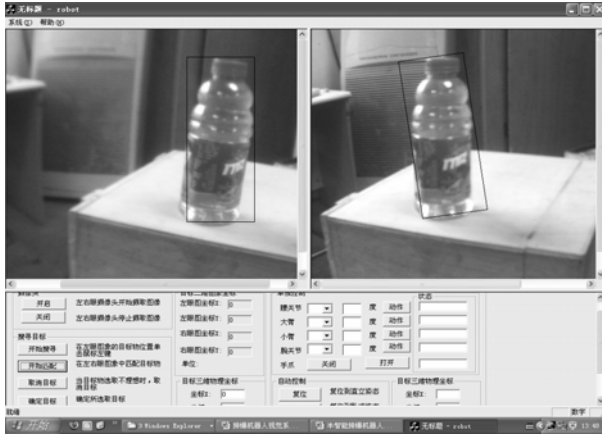


Fig. 5: Human-machine interface which can be used to mark the target.

## 4. Motion control subsystem

### 4.1. Coordinate transformation

The camera-based 3d coordinate acquired in the binocular stereo vision subsystem needs to be transformed into the robot-based 3d coordinate, which is necessary for the motion planning of the manipulator. After the coordinate transformation, the generated robot-based 3d coordinate will be sent to the embedded motion control subsystem for the automatically target positioning of the manipulator.

The coordinate transformation can be divided into two steps: (1) transformation from  $C_{left}$  in the left camera coordinate system to  $P_{claw}$  in the claw coordinate system; (2) translation from  $P_{claw}$  in the claw coordinate system to  $P_{robot}$  in the base coordinate system, see Fig.6, where {Camera}, {Claw}, {Base} represent the left camera coordinate system, the claw coordinate system and the base coordinate system respectively. The left camera coordinate system is centered at the optical center of the left camera, with

the x axis the same as the optical axis; the claw coordinate system is centered at the claw of the manipulator, with axis is parallel to the left camera coordinate system; while the base coordinate system is the basic coordinate system of the robot, see Fig.6.

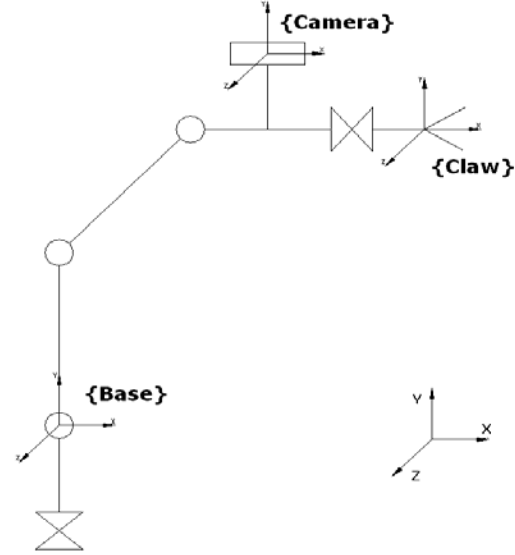


Fig. 6: Coordinate transformation.

#### (1) From $C_{left}$ to $P_{claw}$

The relative location of the claw coordinate system with respect to the left camera coordinate system is invariable, and the corresponding axis is parallel to each other, then the coordinate transformation from  $C_{left}$  to  $P_{claw}$  can be simply achieved by adding a excursion E.

$$P_{claw} = C_{left} + E$$

#### (2) From $P_{claw}$ to $P_{robot}$

The coordinate transformation from  $P_{claw}$  in the claw coordinate system to  $P_{robot}$  in the base coordinate system can be achieved by the kinematic analysis of the manipulator, which is determined by the current pose of the manipulator:

$$\begin{bmatrix} P_{robot} \\ 1 \end{bmatrix} = {}^B T_C \times \begin{bmatrix} P_{claw} \\ 1 \end{bmatrix}$$

Where  ${}^B T_C$  is the homogeneous transform matrix from the claw coordinate system to the base coordinate system, it can be obtained by acquiring the current pose of the manipulator.

### 4.2. Control of manipulator

The embedded motion control subsystem, which is located on the vehicle, is a local control center of the robot. It performs the tasks of receiving commands

and corresponding data from the remote server, controlling the vehicle and the manipulator and reporting the status of the robot to the remote server.

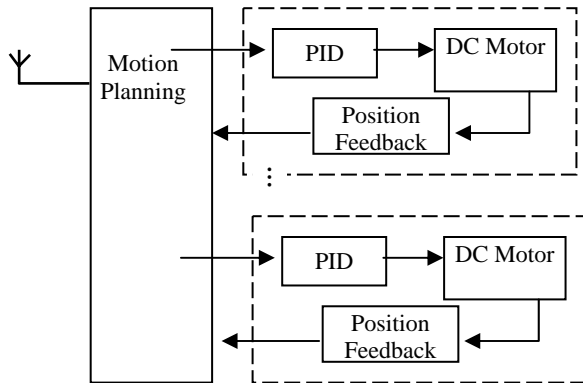


Fig. 7: Embedded motion control subsystem.

In order to perform the task of automatically target positioning, the embedded motion control subsystem receive the robot based 3d coordinate of the target from the remote server, then arrange the motion of the manipulator and control the motors to complete the movement.

The embedded motion control subsystem is implemented under the platform of PC104, where an integrated control system was developed [8]-[13]. The integrated control system will arrange the motion of the manipulator and also control the dc motors of the manipulator, see Fig.7.

## 5. Conclusions

Automatic target positioning is necessary for a bomb-disposal robot to perform the duty of disposing a perilous target. We proposed the design of a bomb-disposal robot based on binocular stereo vision which can achieve the state-of-the-art of automatically target positioning. Through the coordination of the binocular stereo vision subsystem, the coordinate transformation and the embedded motion control subsystem of the whole bomb-disposal robot system, a satisfied level of target positioning accuracy is acquired through the experimental result.

## Acknowledgement

This work is supported by Chinese postdoctoral science funds (Grant No. 20060400217).

## References

- [1] A. R. Graves, C. Czarnecki, Distributed Generic Control for Multiple Types of Telerobot. *ICRA 1999*: 2209-2214, 1998.
- [2] H. Silven, A four-step camera calibration procedure with implicit. image correction. In *Proc. IEEE Conference on Computer Vision and Pattern. Recognition*, 1: 1106-1112, 1997.
- [3] T.A. Clarke and J.G. Fryer, The Development of Camera Calibration Methods and Models. *Photogrammetric Record*, 16: 51-66, 1998.
- [4] J. Weng, P. Cohen and M. Herniou, Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14: 965-980, 1992.
- [5] O. D. Faugeras and G. Toscani, Camera calibration for 3D computer vision. *Proc. International Workshop on Industrial Applications of Machine Vision and Machine Intelligence*, Silken, Japan, pp. 240-247, 1987.
- [6] R. Y. Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, 3: 323-344.
- [7] Z. Zhang, Flexible camera calibration by viewing a plane from unknown orientation. *Proceedings of the 7th IEEE International Conference on Computer Vision*. IEEE Computer Society Press, pp. 666-673, 1999.
- [8] *MATLAB User's Guide*. The Math Works Inc., September 2000.
- [9] *Simulink User's Guide*. The Math Works Inc., September 2000.
- [10] *Real-Time Workshop User's Guide*. The  $M_t^{*--}$  Works Inc., September 2000.
- [11] E. Bianchi and L. Dozio. Some experience in fast hard real-time control in user space with RTAI-LXRT. *Real Time Linux Workshop*, Orlando, FL, 2000.
- [12] N. Costescu, D. Dawson, M. Loffler, QMotor 2.0 – A Real-Time PC Based Control Environment, *IEEE Control Systems Magazine*, pp. 68-76, June 1999.
- [13] V. Yodaiken, M. Barabanov, Real-time Linux, *Linw Journal*, 1997.