

Article

Gaze Point Tracking Based on a Robotic Body–Head–Eye Coordination Method

Xinyang Feng ^{1,†}, Qingbin Wang ^{2,†}, Hua Cong ¹, Yu Zhang ¹ and Mianhao Qiu ^{1,*}

¹ Army Academy of Armored Forces, Beijing 100072, China; m17695720631@163.com (X.F.); 18911025632@163.com (H.C.); zhangyuzgy2018@163.com (Y.Z.)

² Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China; maokang94@163.com

* Correspondence: 15801000128@163.com

† These authors contributed equally to this work.

Abstract: When the magnitude of a gaze is too large, human beings change the orientation of their head or body to assist their eyes in tracking targets because saccade alone is insufficient to keep a target at the center region of the retina. To make a robot gaze at targets rapidly and stably (as a human does), it is necessary to design a body–head–eye coordinated motion control strategy. A robot system equipped with eyes and a head is designed in this paper. Gaze point tracking problems are divided into two sub-problems: in situ gaze point tracking and approaching gaze point tracking. In the in situ gaze tracking state, the desired positions of the eye, head and body are calculated on the basis of minimizing resource consumption and maximizing stability. In the approaching gaze point tracking state, the robot is expected to approach the object at a zero angle. In the process of tracking, the three-dimensional (3D) coordinates of the object are obtained by the bionic eye and then converted to the head coordinate system and the mobile robot coordinate system. The desired positions of the head, eyes and body are obtained according to the object's 3D coordinates. Then, using sophisticated motor control methods, the head, eyes and body are controlled to the desired position. This method avoids the complex process of adjusting control parameters and does not require the design of complex control algorithms. Based on this strategy, in situ gaze point tracking and approaching gaze point tracking experiments are performed by the robot. The experimental results show that body–head–eye coordination gaze point tracking based on the 3D coordinates of an object is feasible. This paper provides a new method that differs from the traditional two-dimensional image-based method for robotic body–head–eye gaze point tracking.

Keywords: bionic eyes; gaze point tracking; gaze point approaching; body–eye–head coordination; 3D coordinates



Citation: Feng, X.; Wang, Q.; Cong, H.; Zhang, Y.; Qiu, M. Gaze Point Tracking Based on a Robotic Body–Head–Eye Coordination Method. *Sensors* **2023**, *23*, 6299. <https://doi.org/10.3390/s23146299>

Academic Editors: David Chenelet and Stephen Monk

Received: 1 June 2023

Revised: 29 June 2023

Accepted: 6 July 2023

Published: 11 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

When the magnitude of a gaze is too large, human beings change the orientation of their head or body to assist their eyes in tracking targets because saccade alone is insufficient to keep a target at the center region of the retina. Studies on body–head–eye coordination gaze point tracking are still rare because the body–head–eye coordination mechanism of humans is prohibitively complex. Multiple researchers have investigated the eye–head coordination mechanism, binocular coordination mechanism and bionic eye movement control. In addition, researchers have validated the eye–head coordination models on eye–head systems. This work is significant for the development of intelligent robots for human–robot interaction. However, most of these methods are based on the principle of neurology, and their further developments and applications may be limited by people's understanding of human processes. However, binocular coordination based on the 3D coordinates of an object is simple and practical, as verified by our previous paper [1].

When the fixation point transfers greatly, the head and eyes should move in coordination to accurately shift the gaze to the target. Multiple studies have built models of eye–head coordination based on the physiological characteristics of humans. For example, Kardamakis A A et al. [2] researched eye–head movement and gaze shifting. The best balance between eye movement speed and the duration time was sought, and the optimal control method was used to minimize the loss of motion. Freedman E G et al. [3] studied the physiological mechanism of coordinated eye–head movement. However, they did not establish an engineering model. Nakashima et al. [4] proposed a method for gaze prediction that combines information on the head direction with a saliency map. In another study [5], the authors presented a robotic head for social robots to attend to scene saliency with bio-inspired saccadic behaviors. The scene saliency was determined by measuring low-level static scene information, motion, and prior object knowledge. Law et al. [6] described a biologically constrained architecture for developmental learning of eye–head gaze control on an iCub robot. They also identified stages in the development of infant gaze control and proposed a framework of artificial constraints to shape the learning of the robot in a similar manner. Other studies have investigated the mechanisms of eye–head movement for robots and achieved satisfactory performance [7,8].

Some application studies based on coordinated eye–head movement have been carried out in addition to the mechanism research. For example, Kuang et al. [9] developed a method for egocentric distance estimation based on the parallax that emerges during compensatory head–eye movements. This method was tested in a robotic platform equipped with an anthropomorphic neck and two binocular pan–tilt units. Reference [10]’s model is capable of reaching static targets posed at a starting distance of 1.2 m in approximately 250 control steps. Hülse et al. [11] introduced a computational framework that integrates robotic active vision and reaching. Essential elements of this framework are sensorimotor mappings that link three different computational domains relating to visual data, gaze control and reaching.

Some researchers have applied the combined movement of the eyes, head and body in mobile robots. In one study [12], large reorientations of the line of sight, involving combined rotations of the eyes, head, trunk and lower extremities, were executed either as fast single-step or as slow multiple-step gaze transfers. Daye et al. [13] proposed a novel approach for the control of linked systems with feedback loops for each part. The proximal parts had separate goals. In addition, an efficient and robust human tracker for a humanoid robot was implemented and experimentally evaluated in another study [14].

On the one hand, human eyes can obtain three-dimensional (3D) information from objects. This 3D information is useful for humans to make decisions. Human can shift their gaze stably and approach a target using the 3D information of the object. When the human gaze shifts to a moving target, the eyes first rotate to the target, and then the head and even the body rotate if the target leaves the sight of the eyes [15]. Therefore, the eyes, head and body move in coordination to shift the gaze to the target with minimal energy expenditure. On the other hand, when a human approaches a target, the eyes, head and body rotate to face the target and the body moves toward the target. The two movements are typically executed with the eyes, head and body acting in conjunction. A robot that can execute these two functions will be more intelligent. Such a robot would need to exploit the smooth pursuit of eyes [16], coordinated eye–head movement [17], target detection and the combined movement of the eyes, head and robot body to carry out these two functions. Studies have achieved many positive results in these aspects.

Mobile robots can track and locate objects according to 3D information. Some special cameras such as deep cameras and 3D lasers have been applied to obtain the 3D information of the environment and target. In one study [18], a nonholonomic under-actuated robot with bounded control was described that travels within a 3D region. A single sensor provided the value of an unknown scalar field at the current location of the robot. Nefti-Meziani S et al. [19] presented the implementation of a stereo-vision system integrated in a humanoid robot. The low cost of the vision system is one of the main

aims, avoiding expensive investment in hardware when used in robotics for 3D perception. Namavari A et al. [20] presented an automatic system for the gauging and digitalization of 3D indoor environments. The configuration consisted of an autonomous mobile robot, a reliable 3D laser rangefinder and three elaborated software modules.

The main forms of motion of bionic eyes include saccade [1], smooth pursuit, vergence [21], vestibule–ocular reflex (VOR) [22] and optokinetic reflex (OKR) [23]. Saccade and smooth pursuit are the two most important functions of the human eye. Saccade is used to move eyes voluntarily from one point to another by rapid jumping, while smooth pursuit can be applied to track moving targets. In addition, binocular coordination and eye–head coordination are of high importance to realize object tracking and gaze control.

It is of great significance for robots to be able change their fixation point quickly. In control models, the saccade control system should be implemented using a position servo controller to change and keep the target at the center region of the retina with minimum time consumption. Researchers have been studying the implementation of saccade on robots over the last twenty years. For example, in 1997, Bruske et al. [24] incorporated saccadic control into a binocular vision system by using the feedback error learning (FEL) strategy. In 2013, Wang et al. [25] designed an active vision system that can imitate saccade and other eye movements. The saccadic movements were implemented with an open-loop controller, which ensures faster saccadic eye movements than a closed-loop controller can accommodate. In 2015, Antonelli et al. [26] achieved saccadic movements on a robot head by using a model called recurrent architecture (RA). In this model, the cerebellum is regarded as an adaptive element used to learn an internal model, while the brainstem is regarded as a fixed-inverse model. The experimental results on the robot showed that this model is more accurate and less sensitive to the choice of the inverse model relative to the FEL model.

The smooth pursuit system acts as a velocity servo controller to rotate eyes at the same angular rate as the target while keeping them oriented toward the desired position or in the desired region. In Robinson's model of smooth pursuit [27], the input is the velocity of the target's image across the retina. The velocity deviation is taken as the major stimulus to pursue and is transformed into an eye velocity command. Based on Robinson's model, Brown [28] added a smooth predictor to accommodate time delays. Deno et al. [29] applied a dynamic neural network, which unified two apparently disparate models of smooth pursuit and dynamic element organization to the smooth pursuit system. The dynamic neural network can compensate for delays from the sensory input to the motor response. Lunghi et al. [30] introduced a neural adaptive predictor that was previously trained to accomplish smooth pursuit. This model can explain a human's ability to compensate for the 130 ms physiological delay when they follow external targets with their eyes. Lee et al. [31] applied a bilateral OCS model on a robot head and established rudimentary prediction mechanisms for both slow and fast phases. Avni et al. [32] presented a framework for visual scanning and target tracking with a set of independent pan–tilt cameras based on model predictive control (MPC). In another study [33], the authors implemented smooth pursuit eye movement with prediction and learning in addition to solving the problem of time delays in the visual pathways. In addition, some saccade and smooth pursuit models have been validated on bionic eye systems [34–37]. Santini F et al. [34] showed that the oculomotor strategies by which humans scan visual scenes produce parallaxes that provide an accurate estimation of distance. Other studies have realized the coordinated control of eye and arm movements through configuration and training [35]. Song Y et al. [36] proposed a binocular control model, which was derived from a neural pathway, for smooth pursuit. In their smooth pursuit experiments, the maximum retinal error was less than 2.2° , which is sufficient to keep a target in the field of view accurately. An autonomous mobile manipulation system was developed in the form of a modified image-based visual servo (IBVS) controller in a study [37].

The above-mentioned work is significant for the development of intelligent robots. However, there are some shortcomings. First, most of the existing methods are based on

the principle of neurology, and further developments and applications may be limited by people's understanding aimed at humans. Second, only two-dimensional (2D) image information is applied when gaze shifts to targets are implemented, while 3D information is ignored. Third, the studies of smooth pursuit [16], eye-head coordination [17], gaze shift and approach are independent and have not been integrated. Fourth, bionic eyes are different from human eyes; for example, some of them are two eyes that are fixed without movement or move with only 1 DOF, whereas some of them use special cameras or a single camera. Fifth, the movements of bionic eyes and heads are performed separately, without coordination.

To overcome the shortcomings mentioned above to a certain extent, a novel control method that implements the gaze shift and approach of a robot according to 3D coordinates is proposed in this paper. A robot system equipped with bionic eyes, a head and a mobile robot is designed to help nurses deliver medicine in hospitals. In this system, both the pan and each eye have 2 DOF (namely, tilt and pan [38]), and the mobile robot can rotate and move forward over the ground. When the robot gaze shifts to the target, the 3D coordinates of the target are acquired by the bionic eyes and transferred to the eye coordination system, head coordination system and robot coordination system. The desired position of the eye, head and robot are calculated based on the 3D information of the target. Then, the eye, head and mobile robot are driven to the desired positions. When the robot approaches the target, the eye, head and mobile robot first rotate to the target and then move to the target. This method allows the robot to achieve the above-mentioned functions with minimal resource consumption and can separate the control of the eye, head and mobile robot, which can improve the interactions between robots, human beings and the environment.

The rest of the paper is organized as follows. In Section 2, the robot system platform is introduced, and the control system is presented. In Section 3, the desired position is discussed and calculated. Robot pose control is described in Section 4. The experimental results are given and discussed in Section 5; finally, conclusions are drawn in Section 6.

2. Platform and Control System

To study the gaze point tracking of the robot, this paper designs a robot experiment platform including the eye-head subsystem and the mobile robot subsystem.

2.1. Robot Platform

The physical object of the robot is shown in Figure 1. With the mobile robot as a carrier, a head with two degrees of freedom is fixed on the mobile robot, and the horizontal and vertical rotations of the head are controlled by M_{hu} and M_{hd} , respectively. The bionic eye system is fixed to the head. The mobile robot is driven by two wheels, each of which is individually controlled by a servo motor. The angle and displacement of the robot platform can be determined by controlling the distance and speed of each wheel's movement. The output shaft of each stepper motor of the head and eye is equipped with a rotary encoder to detect the position of the motor. Using the frequency multiplication technique, the resolution of the rotary encoder is 0.036° . The purpose of using a rotary encoder is to prevent the effects of lost motor motion on the 3D coordinate calculations. The movement of each motor is limited by a limit switch. The initial positioning of the eye system is based on the visual positioning plate [39].

The robot system includes two eyes and one mobile robot. To simulate the eyes and the head, six DOFs are designed in this system. The left eye's pan and tilt are controlled by motors M_{lu} and M_{ld} , respectively. The right eye's pan and tilt are controlled by motors M_{ru} and M_{rd} , respectively. The head's pan and tilt are controlled by motors M_{hu} and M_{hd} , respectively. The mobile robot has two driving wheels and can perform rotation and forward movement. When the mobile robot needs to rotate, two wheels are set to turn the same amount in different directions. When the mobile robot needs to go forward, two wheels are set to turn the same amount in the same direction.

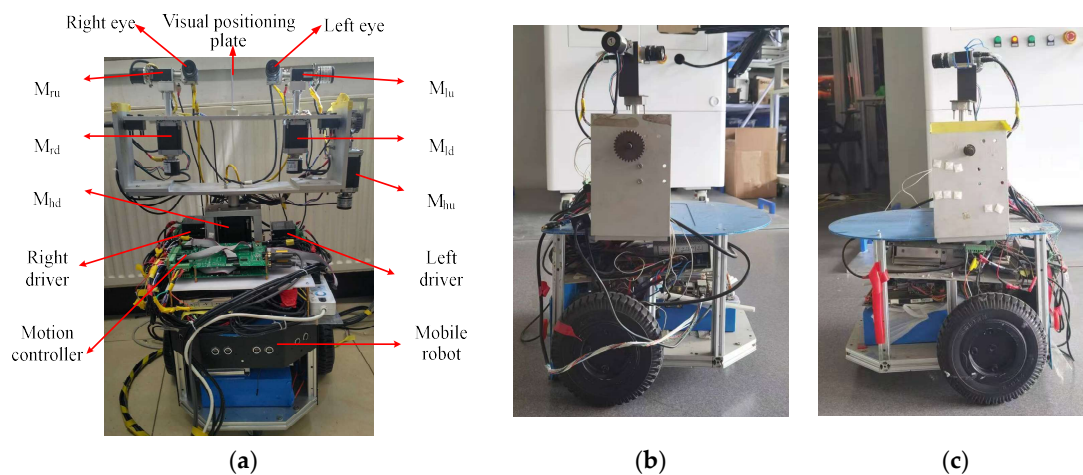


Figure 1. Physical implementation of the robot system. (a) The front side. (b) The left side. (c) The right side.

A diagram of the robot system's organization is shown in Figure 2. The host computer and the mobile robot motion controller, the head motion controller and the eye motion controller all communicate through the serial ports. For satisfactory communication quality and stability, the baud rate of serial communication is 9600 bps. The camera communicates with the host computer via a GigE Gigabit Network. The camera's native resolution is 1600×1200 pixels. To increase the calculation speed, the system uses an image downsampled to 400×300 pixels.

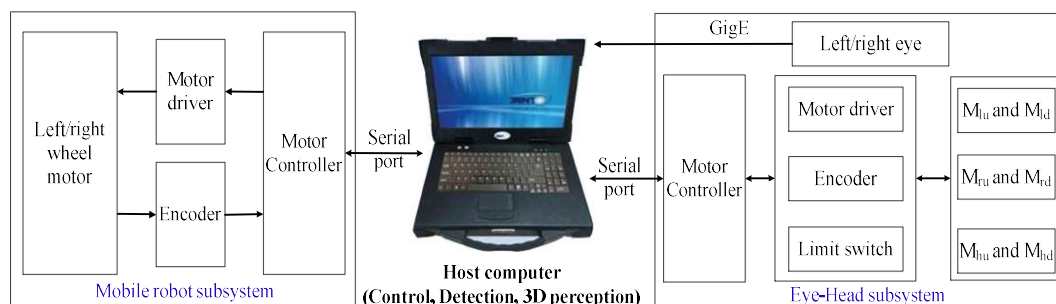


Figure 2. Robot system's organization diagram.

2.2. Control System

Figure 3 shows the control block diagram of the gaze point tracking of the mobile robot. First, based on binocular stereo-vision perception, the binocular pose and the left and right images are used to calculate the 3D coordinates of the target [40], and the coordinates of the target in the eye coordinate system are converted to the head and mobile robot coordinate system. Then, the desired poses of the eyes, head and mobile robot are calculated according to the 3D coordinates of the target. Finally, according to the desired pose, the motor is controlled to move to the desired position, and the change in the position of the motor is converted into changes in the eyes, head and mobile robot.

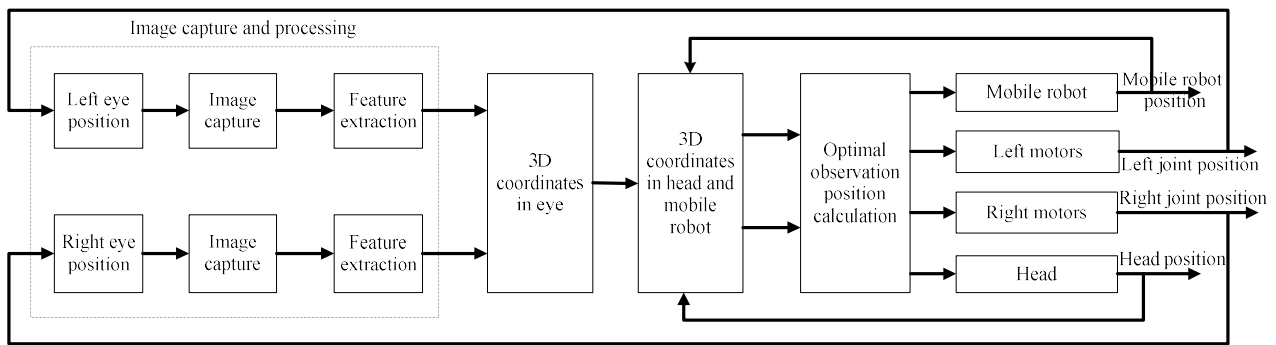


Figure 3. Block diagram of the gaze point tracking control system.

The tracking and approaching motion control problem based on the target 3D coordinates [1] is equivalent to solving the index J minimization problem of Equation (1), where f_i is the current state vector of the joint pose of the eye, head and mobile robot and f_q is the desired state vector:

$$J = \|f_i - f_q\| \quad (1)$$

where J is the indicator function.

Figure 4a shows the definition of each coordinate system of the robot. The coordinate system of the eye is $O_e X_e Y_e Z_e$, which coincides with the left motion module's base coordinate system at the initial position. The head coordinate system is $O_h X_h Y_h Z_h$, and the coordinates $P_h(x_h, y_h, z_h)$ of the point P in the head coordinate system can be calculated using the coordinates $P_e(x_e, y_e, z_e)$ in the eye coordinate system. The definitions of d_x and d_y are shown in Figure 4b. The robot coordinate system $O_w X_w Y_w Z_w$ coincides with the head coordinate system of the initial position. In the bionic eye system, the axis of rotation of the robot approximately coincides with Y_w .

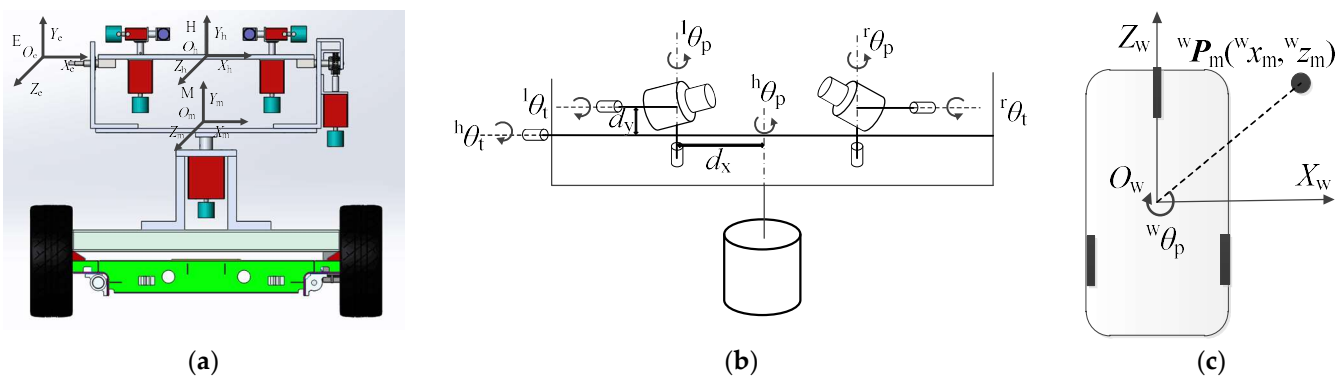


Figure 4. Robot coordinates system and system parameter definition, (a) coordinate system definition, (b) eye-head system parameters and (c) mobile robot parameters.

Figure 4b,c show the definition of each system parameter. $l\theta_p$ and $l\theta_t$ are the pan and tilt of the left eye, respectively. $r\theta_p$ and $r\theta_t$ are the pan and tilt of the right eye, respectively. $h\theta_p$ and $h\theta_t$ are the pan and tilt of the head, respectively. The angle of the robot that rotates around the Y_w axis is $w\theta_p$. The robot can not only rotate around Y_w but can also shift in the $X_w O_w Z_w$ plane. When the robot moves, the robot coordinate system at time i is the base coordinate system, and the position of the robot at time $i + 1$ relative to the base coordinate system is ${}^w P_m ({}^w x_m, {}^w z_m)$. When the robot performs gaze point tracking or approaches the target, the 3D coordinates of the target are first calculated at time i , and then the desired posture f_q of each part of the robot at time $i + 1$ is calculated according to the 3D coordinates of the target. When the current pose f_i of the robot system is equal to the desired pose, the robot maintains the current pose; when not equal, the system controls the various parts

of the robot to move to the desired pose. The current pose vector of the robot system is $f_i = ({}^w x_{mi}, {}^w z_{mi}, {}^w \theta_{pi}, {}^h \theta_{pi}, {}^h \theta_{ti}, {}^l \theta_{pi}, {}^l \theta_{ti}, {}^r \theta_{pi}, {}^r \theta_{ti})$, and the desired pose is $f_q = ({}^w x_{mq}, {}^w z_{mq}, {}^w \theta_{pq}, {}^h \theta_{pq}, {}^h \theta_{tq}, {}^l \theta_{pq}, {}^l \theta_{tq}, {}^r \theta_{pq}, {}^r \theta_{tq})$. When performing in situ gaze point tracking, the robot performs only pure rotation and does not move forward. When the robot approaches the target, it first turns to the target and then moves straight toward the target. Therefore, the definition of f_q in the two tasks is different. Let ${}^s f_q$ be the desired pose when the gaze point is tracked and ${}^a f_q$ be the desired pose of the robot when approaching the target.

After analyzing the control system, we found that the most important step in solving this control problem is to determine the desired pose.

3. Desired Pose Calculation

When performing in situ gaze point tracking, the robot performs only pure rotation and does not move forward. When the robot approaches the target, it first turns to the target and then moves straight toward the target. Therefore, the calculation of the desired pose can be divided into two sub-problems: (1) desired pose calculation for in situ gaze point tracking and (2) desired pose calculation for approaching gaze point tracking.

The optimal observation position is used for the accurate acquisition of 3D coordinates. The 3D coordinate accuracy is related to the baseline, time difference and image distortion. In the bionic eye platform, the baseline is changed with the changes in the cameras' positions because the optical center is not coincident with the center of rotation. The 3D coordinate error of the target is smaller when the baseline of the two cameras is longer. Therefore, it is necessary to keep the baseline unchanged. On the other hand, there is a time difference caused by unstick synchronization between image acquisition and camera position acquisition. In addition, it is necessary to keep the target in the center areas of the two camera images to obtain accurate 3D coordinates of the target.

3.1. Optimal Observation Position of Eyes

In the desired pose of the robot, the most important aspect is the expected pose of the bionic eye [40]. Following the definition of this parameter, the calculation of the desired pose of the robot system is greatly simplified; thus, we present an engineering definition here of the desired pose of the bionic eye.

As shown in Figure 5, ${}^l m_i ({}^l u_i, {}^l v_i)$ and ${}^r ({}^r u_i, {}^r v_i)$ are the image coordinates of point ${}^e P$ in the camera at time i . ${}^l m_o$ and ${}^r m_o$ are the image centers of the left and right cameras, respectively. ${}^l P$ is the vertical point of ${}^e P$ along the line ${}^l O_c {}^l Z_c$, and ${}^r P$ is the vertical point of ${}^e P$ along the line ${}^r O_c {}^r Z_c$. ${}^l \Delta m$ is the distance between ${}^l m$ and ${}^l m_o$. ${}^r \Delta m$ is the distance between ${}^r m$ and ${}^r m_o$. D_b is the baseline length. The pan angles of the left and right cameras in the optimal observation position are ${}^l \theta_p$ and ${}^r \theta_p$, respectively. The tilt angles of the left and right cameras in the optimal observation position are ${}^l \theta_t$ and ${}^r \theta_t$, respectively. $P_{ob} ({}^l \theta_p, {}^l \theta_t, {}^r \theta_p, {}^r \theta_t)$ is the optimal observation position.

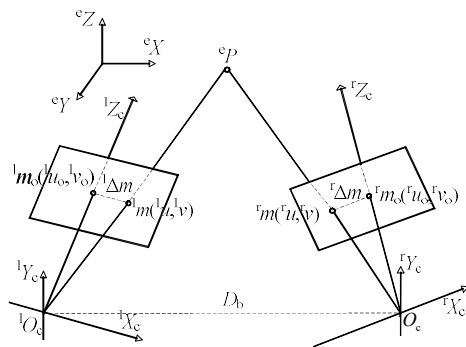


Figure 5. Schematic of the relationship between a Cartesian point and its image point.

When the two eyeballs of the bionic eye move relative to each other, the 3D coordinates of the target obtained by the bionic eye produces a large error. To characterize this error,

we give a detailed analysis of its origins in Appendix A. Through analysis, we obtain the following conclusions to reduce the measurement error of the bionic eye:

- (1) Make the length of D_b long enough, and maintain as much length as possible during the movement;
- (2) Try to observe the target closer to the target so that the depth error is as small as possible;
- (3) During the movement of the bionic eye, control the two cameras so that they move at the same angular velocity;
- (4) Try to keep the target symmetrical, and make ${}^l\Delta m$ and ${}^r\Delta m$ as equal as possible in the left and right camera images.

Based on these four methods, the motion strategy of the motor is designed, and the measurement accuracy of the target's 3D information can be effectively improved.

According to the conclusion, we can define a definition of the optimal observed pose of the bionic eye to reduce the measurement error.

The optimal observation position needed to meet the conditions is listed in Equation (2). When the target is very close to the eyes, the target's optimal observation position cannot be obtained because the image position of the target can be kept at the image center region. It is challenging to obtain the optimal solution of the observation position based on Equation (12). However, a suboptimal solution can be obtained by using a simplified calculation method. First, ${}^l\theta_t$ and ${}^r\theta_t$ are calculated in the case that ${}^l\theta_t$ and ${}^r\theta_t$ are equal to zero; then, ${}^l\theta_t$ and ${}^r\theta_t$ are calculated while ${}^l\theta_t$ and ${}^r\theta_t$ are kept equal to the calculated value. Trial-and-error methods can be used to obtain the optimal solution when the suboptimal solution is obtained.

$$\begin{cases} {}^l\theta_{pq} = {}^r\theta_{pq} = \theta_p \\ {}^l\theta_{tq} = {}^r\theta_{tq} = \theta_t \\ {}^l\Delta m = -{}^r\Delta m \end{cases} \quad (2)$$

where

$${}^l\Delta m = \begin{pmatrix} {}^l\Delta u \\ {}^l\Delta v \end{pmatrix} = \begin{pmatrix} {}^l u_i - {}^l u_0 \\ {}^l v_i - {}^l v_0 \end{pmatrix} \quad (3)$$

$${}^r\Delta m = \begin{pmatrix} {}^r\Delta u \\ {}^r\Delta v \end{pmatrix} = \begin{pmatrix} {}^r u_i - {}^r u_0 \\ {}^r v_i - {}^r v_0 \end{pmatrix} \quad (4)$$

3.2. Desired Pose Calculation for In Situ Gaze Point Tracking

When the range of target motion is large and the desired posture of the eyeball exceeds its reachable posture, the head and mobile robot move to keep the target in the center region of the image. In robotic systems, eye movements tend to consume the least amount of resources and do not have much impact on the stability of the head and mobile robot during exercise. Head rotation consumes more resources than the eyeball but consumes fewer resources than trunk rotation. At the same time, the rotation of the head affects the stability of the eyeball but does not have much impact on the stability of the trunk. Mobile robot rotation consumes the most resources and has a large impact on the stability of the head and eyeball. When tracking the target, one needs only to keep the target in the center region of the binocular image. Therefore, when performing gaze point tracking, the movement mechanism of the head, eyes and mobile robot are designed with the principle of minimal resource consumption and maximum system stability. When the eyeball can perceive the 3D coordinates of the target in the reachable and optimal viewing posture, only the eye is rotated; otherwise, the head is rotated. The head also has an attainable range of poses. When the desired pose exceeds this range, the mobile robot needs to be turned so that the bionic eye always perceives the 3D coordinates of the target in the optimal viewing position. Let ${}^h\gamma_p$ and ${}^h\gamma_t$ be the angles between the head and the gaze point in the $X_h O_h Z_h$ and $Y_h O_h Z_h$ planes, respectively. The range of binocular rotation in the horizontal direction is $[-{}^e\theta_{pmax}, {}^e\theta_{pmax}]$, and the range of binocular rotation in the vertical direction is $[-{}^e\theta_{tmax}, {}^e\theta_{tmax}]$. The range of head rotation in the horizontal direction is $[-{}^h\theta_{pmax},$

${}^h\theta_{pmax}$], and the range of head rotation in the vertical direction is $[-{}^h\theta_{tmax}, {}^h\theta_{tmax}]$. For the convenience of calculation, the angles between the head and the fixation point in the horizontal direction and the vertical direction are designated as $[-{}^h\gamma_{pmax}, {}^h\gamma_{pmax}]$ and $[-{}^h\gamma_{tmax}, {}^h\gamma_{tmax}]$, respectively. When the angle between the head and the target exceeds a set threshold, the head needs to be rotated to the ${}^h\theta'_p$ and ${}^h\theta'_t$ positions in the horizontal and vertical directions, respectively. When ${}^h\theta'_p$ exceeds the angle that the head can attain, the angle at which the mobile robot needs to be compensated is ${}^w\theta_p$. In the in situ gaze point tracking task, the cart does not need to translate in the $X_wO_wZ_w$ plane, so $x_w = 0$, and $z_w = 0$. Furthermore, according to the definition of the optimal observation pose of the bionic eye, the conditions that ${}^g f_q$ should satisfy are

$${}^g f_q = \begin{pmatrix} {}^w x_{mq} = 0 \\ {}^w z_{mq} = 0 \\ {}^w \theta_{pq} = \{ \theta \mid |\theta| \leq 2\pi, {}^h\theta_{pq} + \theta = {}^h\theta'_p \} \\ {}^h\theta_{pq} = \{ \theta \mid |\theta| \leq {}^h\theta_{pmax}, |{}^h\gamma_p| \leq {}^h\gamma_{pmax} \} \\ {}^h\theta_{tq} = \{ \theta \mid |\theta| \leq {}^h\theta_{tmax}, |{}^h\gamma_t| \leq {}^h\gamma_{tmax} \} \\ {}^1\theta_{pq} = {}^r\theta_{pq} = \{ \theta \mid |\theta| \leq {}^e\theta_{pmax}, {}^1\Delta m_l = -\Delta m_r \} \\ {}^1\theta_{tq} = {}^r\theta_{tq} = \{ \theta \mid |\theta| \leq {}^e\theta_{tmax}, \Delta m_l = -\Delta m_r \} \end{pmatrix} \tag{5}$$

The desired pose needs to be calculated based on the 3D coordinates of the target. Therefore, to obtain the desired pose, it is necessary to acquire the 3D coordinates of the target according to the current pose of the robot.

3.2.1. Three-Dimensional Coordinate Calculation

The mechanical structure and coordinate settings of the system are shown in Figure 6a. The principle of binocular stereoscopic 3D perception is shown in Figure 6b. E is the eye coordinate system, E_l is the left motion module's end coordinate system, E_r is the right motion module's end coordinate system, B_l is the left motion module's base coordinate system, B_r is the right motion module's base coordinate system, C_l is the left camera coordinate system and C_r is the right camera coordinate system. In the initial position, E_l coincides with B_l , and E_r overlaps with B_r . When the binocular system moves, the base coordinate system does not change. 1T represents the transformation matrix of the eye coordinate system E to the left motion module's base coordinate system B_l , rT represents the transformation matrix of E to B_r , 1T_e represents the transformation matrix of B_l to E_l , rT_e represents the transformation matrix of B_r to E_r and 1T_m represents the leftward motion. The module end coordinate system corresponds to the transformation matrix of the left camera coordinate system, and rT_m represents the transformation matrix of the right motion module's end coordinate system to the right camera coordinate system. 1T_r represents the transformation matrix of the right camera coordinate system to the left camera coordinate system at the initial position.

The origin 1O_c of C_l lies at the optical center of the left camera, the 1Z_c axis points in the direction of the object parallel to the optical axis of the camera, the 1X_c axis points horizontally to the right along the image plane and the 1Y_c axis points vertically downward along the image plane. The origin rO_c of C_r lies at the optical center of the right camera, rZ_c is aligned with the direction of the object parallel to the optical axis of the camera, rX_c points horizontally to the right along the image plane and rY_c points vertically downward along the image plane. E_l 's origin 1O_e is set at the intersection of the two rotation axes of the left motion module, 1Z_e is perpendicular to the two rotation axes and points to the front of the platform, 1X_e coincides with the vertical rotation axis and 1Y_e coincides with the horizontal rotation axis. Similarly, the origin rO_e of the coordinate system E_r is set at the intersection of the two rotation axes of the right motion module, rZ_e is perpendicular to the two rotation axes and points toward the front of the platform, rX_e coincides with the vertical rotation axis and rY_e coincides with the horizontal rotation axis.

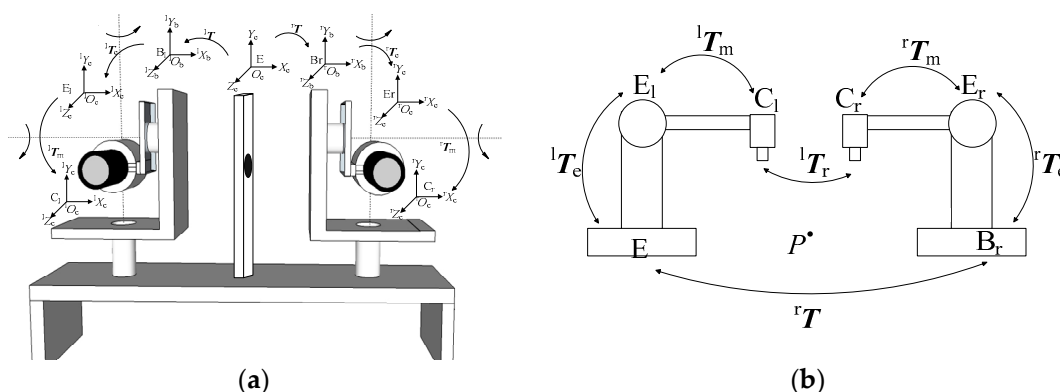


Figure 6. (a) Mechanical structure and coordinate systems of the bionic eye platform and (b) binocular 3D perception principle of bionic eyes.

The left motion module’s base coordinates system B_1 coincides with the eye coordinate system E ; thus, 1T consists of an identity matrix. To calculate the 3D coordinates of the feature points in real time from the camera pose, it is necessary to calculate rT . At the initial position of the system, the external parameters 1T_r of the left and right cameras are calibrated offline, as are the hand–eye parameters of the left–right motion module to the camera coordinate system.

When the system is in its initial configuration, the coordinates of point P in the eye coordinate system are $P_e (x_e, y_e, z_e)$. Its coordinates in B_1 are ${}^1P_e ({}^1x_e, {}^1y_e, {}^1z_e)$, and its coordinates ${}^1P_c ({}^1x_c, {}^1y_c, {}^1z_c)$ in C_1 are

$${}^1P_c = {}^1T_m^{-1}P_e \tag{6}$$

The coordinates ${}^rP_e ({}^rx_e, {}^ry_e, {}^rz_e)$ of point P in B_r are

$${}^rP_e = {}^rTP_e \tag{7}$$

The coordinates ${}^rP_c ({}^rx_c, {}^ry_c, {}^rz_c)$ of point P in C_r are

$${}^rP_c = {}^rT_m^{-1}{}^rTP_e \tag{8}$$

The point in C_r is transformed into C_1 :

$${}^1P_c = {}^1T_r {}^rT_m^{-1}{}^rTP_e \tag{9}$$

Based on the Equations (6) and (9), rT is available:

$${}^rT = {}^rT_m {}^1T_r^{-1}{}^1T_m^{-1} \tag{10}$$

During the movement of the system, when the left motion module rotates by ${}^1\theta_p$ and ${}^1\theta_t$ in the horizontal and vertical directions, respectively, the transformation relationship between B_1 and E_1 is

$${}^1T_e = \begin{pmatrix} \text{Rot}(Y, {}^1\theta_p)\text{Rot}(X, {}^1\theta_t) & 0 \\ 0 & 1 \end{pmatrix} \tag{11}$$

The coordinates of point P in C_1 are

$${}^1P_e = {}^1T_m^{-1}{}^1T_e P_w = {}^1T_d P_e \tag{12}$$

Assume that

$${}^1T_d = \begin{pmatrix} {}^1n_x & {}^1o_x & {}^1a_x & {}^1p_x \\ {}^1n_y & {}^1o_y & {}^1a_y & {}^1p_y \\ {}^1n_z & {}^1o_z & {}^1a_z & {}^1p_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (13)$$

The point ${}^1P_{1c}$ (${}^1x_{1c}$, ${}^1y_{1c}$) at which line P^1O_c intersects ${}^1Z_c = 1$ is

$$\begin{pmatrix} {}^1x_{1c} \\ {}^1y_{1c} \end{pmatrix} = \begin{pmatrix} \frac{{}^1n_x x_e + {}^1o_x y_e + {}^1a_x z_e + {}^1p_x}{{}^1n_z x_e + {}^1o_z y_e + {}^1a_z z_e + {}^1p_z} \\ \frac{{}^1n_y x_e + {}^1o_y y_e + {}^1a_y z_e + {}^1p_y}{{}^1n_z x_e + {}^1o_z y_e + {}^1a_z z_e + {}^1p_z} \end{pmatrix} \quad (14)$$

The image coordinates of ${}^1P_{1c}$ in the left camera are m_l (u_l , v_l), (${}^1x_{1c}$, ${}^1y_{1c}$) and (u_l , v_l) and can be converted by the parameters of the camera. According to the camera's internal parameter model, the following can be obtained:

$$\begin{pmatrix} {}^1x_{1c} \\ {}^1y_{1c} \\ 1 \end{pmatrix} = {}^1M_{in}^{-1} \begin{pmatrix} u_l \\ v_l \\ 1 \end{pmatrix} \quad (15)$$

where ${}^1M_{in}$ is the internal parameter matrix of the left camera. The value of (${}^1x_{1c}$, ${}^1y_{1c}$) can be obtained by the image coordinates of ${}^1P_{1c}$, and the parameters of the left camera can be obtained by substituting (15) into (14):

$$\begin{cases} ({}^1n_x - {}^1x_{1c} {}^1n_z)x_e + ({}^1o_x - {}^1x_{1c} {}^1o_z)y_e + ({}^1a_x - {}^1x_{1c} {}^1a_z)z_e + {}^1p_x - {}^1x_{1c} {}^1p_z = 0 \\ ({}^1n_y - {}^1y_{1c} {}^1n_z)x_e + ({}^1o_y - {}^1y_{1c} {}^1o_z)y_e + ({}^1a_y - {}^1y_{1c} {}^1a_z)z_e + {}^1p_y - {}^1y_{1c} {}^1p_z = 0 \end{cases} \quad (16)$$

During the motion of the system, when the right motion module rotates through ${}^r\theta_p$ and ${}^r\theta_t$ in the horizontal and vertical directions, respectively, the transformation relationship between B_r and E_r is

$${}^rT_e = \begin{pmatrix} \text{Rot}(Y, {}^r\theta_p)\text{Rot}(X, {}^r\theta_t) & 0 \\ 0 & 1 \end{pmatrix} \quad (17)$$

The coordinates of point P in C_r are

$${}^rP_e = {}^rT_m^{-1} {}^rT_e {}^rTP_e = {}^rT_d P_e \quad (18)$$

Assume that

$${}^rT_d = \begin{pmatrix} {}^rn_x & {}^ro_x & {}^ra_x & {}^rp_x \\ {}^rn_y & {}^ro_y & {}^ra_y & {}^rp_y \\ {}^rn_z & {}^ro_z & {}^ra_z & {}^rp_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (19)$$

The point ${}^rP_{1c}$ (${}^rx_{1c}$, ${}^ry_{1c}$) at which line P^rO_c intersects ${}^rZ_c = 1$ is

$$\begin{pmatrix} {}^rx_{1c} \\ {}^ry_{1c} \end{pmatrix} = \begin{pmatrix} \frac{{}^rn_x x_e + {}^ro_x y_e + {}^ra_x z_e + {}^rp_x}{{}^rn_z x_e + {}^ro_z y_e + {}^ra_z z_e + {}^rp_z} \\ \frac{{}^rn_y x_e + {}^ro_y y_e + {}^ra_y z_e + {}^rp_y}{{}^rn_z x_e + {}^ro_z y_e + {}^ra_z z_e + {}^rp_z} \end{pmatrix} \quad (20)$$

The image coordinates of ${}^rP_{1c}$ in the camera, namely, m_r (u_r , v_r), (${}^rx_{1c}$, ${}^ry_{1c}$) and (u_r , v_r), can be converted using the parameters of the camera. According to the camera's internal parameter model, the following can be obtained:

$$\begin{pmatrix} {}^rx_{1c} \\ {}^ry_{1c} \\ 1 \end{pmatrix} = {}^rM_{in}^{-1} \begin{pmatrix} u_r \\ v_r \\ 1 \end{pmatrix} \quad (21)$$

where ${}^rM_{in}$ is the inner parameter matrix of the right camera. The value of $({}^rx_{1c}, {}^ry_{1c})$ can be obtained by the image coordinates of ${}^rP_{1c}$ and the parameters in the camera, and the following can be obtained by substituting (21) into (20):

$$\begin{cases} ({}^rn_x - {}^rx_{1c}{}^rn_z)x_e + ({}^ro_x - {}^rx_{1c}{}^ro_z)y_e + ({}^ra_x - {}^rx_{1c}{}^ra_z)z_e + {}^rp_x - {}^rx_{1c}{}^rp_z = 0 \\ ({}^rn_y - {}^ry_{1c}{}^rn_z)x_e + ({}^ro_y - {}^ry_{1c}{}^ro_z)y_e + ({}^ra_y - {}^ry_{1c}{}^ra_z)z_e + {}^rp_y - {}^ry_{1c}{}^rp_z = 0 \end{cases} \quad (22)$$

Four equations can be obtained from Equations (16) and (22) for x_e, y_e and z_e , and the 3D coordinates of point P_e can be calculated by the least squares method.

The 3D coordinates $P_h(x_h, y_h, z_h)$ in the head coordinate system can be obtained by Equation (23). d_x and d_y are illustrated in Figure 4.

$$\begin{pmatrix} x_h \\ y_h \\ z_h \end{pmatrix} = \begin{pmatrix} x_e - d_x \\ y_e - d_y \\ z_e \end{pmatrix} \quad (23)$$

Let the angles at which the current moment of the head rotate relative to the initial position be ${}^h\theta_{pi}$ and ${}^h\theta_{ti}$; the coordinates of the target in the robot coordinate system are

$$\begin{pmatrix} {}^wx_m \\ {}^wy_m \\ {}^wz_m \\ 1 \end{pmatrix} = \begin{pmatrix} \text{Rot}(X, {}^h\theta_{ti})\text{Rot}(Y, {}^h\theta_{pi}) & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} x_h \\ y_h \\ z_h \\ 1 \end{pmatrix} \quad (24)$$

According to the 3D coordinates of the target in the head coordinate system, the angle between the target and Z_h in the horizontal direction and the vertical direction can be obtained as follows:

$${}^h\gamma_p = \arctan\left(\frac{x_h}{z_h}\right) \quad (25)$$

$${}^h\gamma_t = \arctan\left(\frac{y_h}{z_h}\right) \quad (26)$$

When ${}^h\gamma_p$ and ${}^h\gamma_t$ exceed a set threshold, the head needs to rotate. To leave a certain margin for the rotation of the eyeball and for the convenience of calculation, the angles required for the head to rotate in the horizontal direction and the vertical direction are calculated by the principle shown in Figure 7a,b, respectively. Figure 7a shows the calculation principle of the horizontal direction angle when the target's x coordinates of the head coordinate system is greater than zero. After the head is rotated to ${}^h\theta'_p$, the target point is on the 1Z_e axis of the left motion module end coordinate system, and the left motion module reaches the maximum rotatable threshold ${}^e\theta_{pmax}$. Figure 7b shows the calculation principle of the vertical direction when the target's y coordinates of the head coordinate system are greater than d_y . After the head is rotated to ${}^h\theta'_t$, the target point is on the Z_e axis of the eye coordinate system, and the eye reaches the maximum threshold ${}^e\theta_{tmax}$ that can be rotated.

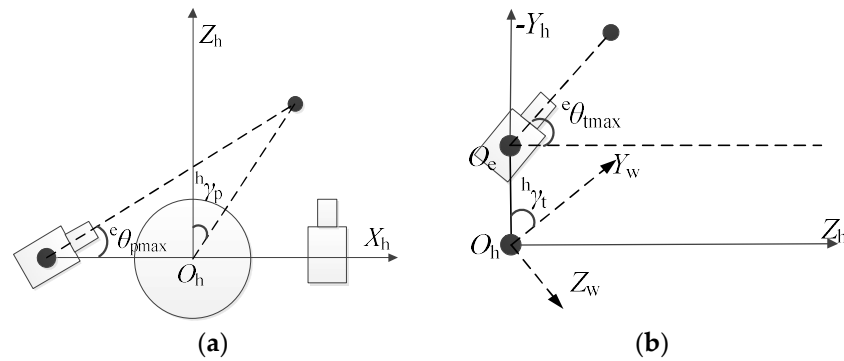


Figure 7. Principle of head rotation calculation in fixation point tracking: (a) horizontal rotation angle and (b) vertical rotation angle.

3.2.2. Horizontal Rotation Angle Calculation

Let the current angle of the head in the horizontal direction be ${}^h\theta_{pi}$. When the head is rotated in the horizontal direction to ${}^h\theta'_p$, the 3D coordinates of the target in the new head coordinate system are

$$\begin{pmatrix} x'_h \\ y'_h \\ z'_h \\ 1 \end{pmatrix} = \begin{pmatrix} \cos({}^h\theta'_p - {}^h\theta_{pi}) & 0 & -\sin({}^h\theta'_p - {}^h\theta_{pi}) & 0 \\ 0 & 1 & 0 & 0 \\ \sin({}^h\theta'_p - {}^h\theta_{pi}) & 0 & \cos({}^h\theta'_p - {}^h\theta_{pi}) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_h \\ y_h \\ z_h \\ 1 \end{pmatrix} \quad (27)$$

Therefore,

$$\begin{pmatrix} x'_h \\ y'_h \\ z'_h \end{pmatrix} = \begin{pmatrix} x_h \cos({}^h\theta'_p - {}^h\theta_{pi}) - z_h \sin({}^h\theta'_p - {}^h\theta_{pi}) \\ y_h \\ x_h \sin({}^h\theta'_p - {}^h\theta_{pi}) + z_h \cos({}^h\theta'_p - {}^h\theta_{pi}) \end{pmatrix} \quad (28)$$

The coordinates of the target in the new eye coordinate system are

$$\begin{pmatrix} {}^e x'_h \\ {}^e y'_h \\ {}^e z'_h \end{pmatrix} = \begin{pmatrix} d_x + x_h \cos({}^h\theta'_p - {}^h\theta_{pi}) - z_h \sin({}^h\theta'_p - {}^h\theta_{pi}) \\ y_h + d_y \\ x_h \sin({}^h\theta'_p - {}^h\theta_{pi}) + z_h \cos({}^h\theta'_p - {}^h\theta_{pi}) \end{pmatrix} \quad (29)$$

After turning, the left motion module reaches the maximum threshold ${}^e\theta_{pmax}$ that can be rotated, so that

$$\tan({}^e\theta_{pmax}) = \frac{{}^e z'_h}{{}^e x'_h} = \frac{x_h \sin({}^h\theta'_p - {}^h\theta_{pi}) + z_h \cos({}^h\theta'_p - {}^h\theta_{pi})}{d_x + x_h \cos({}^h\theta'_p - {}^h\theta_{pi}) - z_h \sin({}^h\theta'_p - {}^h\theta_{pi})} \quad (30)$$

Simplifying Equation (30), we have

$$\sin({}^h\theta'_p - {}^h\theta_{pi}) = \frac{d_x \tan({}^e\theta_{qmax})}{x_h + z_h \tan({}^e\theta_{qmax})} + \frac{x_h \tan({}^e\theta_{qmax}) - z_h}{x_h + z_h \tan({}^e\theta_{qmax})} \cos({}^h\theta'_p - {}^h\theta_{pi}) \quad (31)$$

Assume that

$$\begin{cases} k_1 = \frac{x_h \tan({}^e\theta_{qmax}) - z_h}{x_h + z_h \tan({}^e\theta_{qmax})} \\ k_2 = \frac{d_x \tan({}^e\theta_{qmax})}{x_h + z_h \tan({}^e\theta_{qmax})} \end{cases} \quad (32)$$

According to the triangular relationship,

$$[k_1 \cos({}^h\theta'_p - {}^h\theta_{pi}) + k_2]^2 + \cos^2({}^h\theta'_p - {}^h\theta_{pi}) = 1 \quad (33)$$

The solution of Equation (33) is

$$\cos({}^h\theta'_p - {}^h\theta_{pi}) = \frac{-k_1k_2 \pm \sqrt{k_1^2 - k_2^2 + 1}}{k_1^2 + 1} \quad (34)$$

Therefore,

$${}^h\theta'_p = {}^h\theta_{pi} + \arccos\left(\frac{-k_1k_2 \pm \sqrt{k_1^2 - k_2^2 + 1}}{k_1^2 + 1}\right) \quad (35)$$

Equation (35) has two solutions; therefore, we choose the solution in which the deviation e of Equation (36) is minimized:

$$e = \left| \tan({}^e\theta_{qmax}) - \frac{x_h \sin({}^h\theta'_p - {}^h\theta_{pi}) + z_h \cos({}^h\theta'_p - {}^h\theta_{pi})}{d_x + x_h \cos({}^h\theta'_p - {}^h\theta_{pi}) - z_h \sin({}^h\theta'_p - {}^h\theta_{pi})} \right| \quad (36)$$

When the obtained ${}^h\theta'_p$ is outside of the range $[-{}^h\theta_{pmax}, {}^h\theta_{pmax}]$, the value of ${}^h\theta_{pq}$ is

$${}^h\theta_{pq} = \begin{cases} {}^h\theta_{pmax} & , \quad {}^h\theta'_p \geq {}^h\theta_{pmax} \\ -{}^h\theta_{pmax} & , \quad {}^h\theta'_p \leq -{}^h\theta_{pmax} \\ {}^h\theta'_p & , \quad \text{else} \end{cases} \quad (37)$$

Finally, one can obtain the ${}^w\theta_{pq}$ value:

$${}^w\theta_{pq} = \begin{cases} {}^h\theta'_p - {}^h\theta_{pmax} & , \quad {}^h\theta'_p > {}^h\theta_{pmax} \\ {}^h\theta'_p + {}^h\theta_{pmax} & , \quad {}^h\theta'_p < -{}^h\theta_{pmax} \\ 0 & , \quad \text{else} \end{cases} \quad (38)$$

Based on the same principle, when the x coordinate of the target in the head coordinate system is less than 0, the coordinates of the target in the right motion module base coordinate system after the rotation are

$$\begin{pmatrix} {}^r x'_e \\ {}^r y'_e \\ {}^r z'_e \\ 1 \end{pmatrix} = \begin{pmatrix} x_h \cos({}^h\theta'_p - {}^h\theta_{pi}) - z_h \sin({}^h\theta'_p - {}^h\theta_{pi}) - d_x \\ y_h + d_y \\ x_h \sin({}^h\theta'_p - {}^h\theta_{pi}) + z_h \cos({}^h\theta'_p - {}^h\theta_{pi}) \\ 1 \end{pmatrix} \quad (39)$$

After turning, the right motion module reaches $-{}^e\theta_{pmax}$, and the following can be obtained:

$$\tan(-{}^e\theta_{qmax}) = \frac{{}^r z'_e}{{}^r x'_e} = \frac{x_h \sin({}^h\theta'_p - {}^h\theta_{pi}) + z_h \cos({}^h\theta'_p - {}^h\theta_{pi})}{x_h \cos({}^h\theta'_p - {}^h\theta_{pi}) - z_h \sin({}^h\theta'_p - {}^h\theta_{pi}) - d_x} \quad (40)$$

We simplify Equation (40) as follows:

$$\sin({}^h\theta'_p - {}^h\theta_{pi}) = \frac{d_x \tan({}^e\theta_{qmax})}{x_h - z_h \tan({}^e\theta_{qmax})} - \frac{x_h \tan({}^e\theta_{qmax}) + z_h}{x_h - z_h \tan({}^e\theta_{qmax})} \cos({}^h\theta'_p - {}^h\theta_{pi}) \quad (41)$$

Let

$$\begin{cases} k'_1 = -\frac{x_h \tan({}^e\theta_{qmax}) + z_h}{x_h - z_h \tan({}^e\theta_{qmax})} \\ k'_2 = \frac{d_x \tan({}^e\theta_{qmax})}{x_h - z_h \tan({}^e\theta_{qmax})} \end{cases} \quad (42)$$

The same two solutions are available:

$${}^h\theta'_p = {}^h\theta_{pi} + \arccos\left(\frac{-k'_1k'_2 \pm \sqrt{(k'_1)^2 - (k'_2)^2 + 1}}{(k'_1)^2 + 1}\right) \quad (43)$$

Select the solution in which the deviation e of Equation (44) is minimized:

$$e = \left| -\tan({}^e\theta_{qmax}) - \frac{x_h \sin({}^h\theta'_p - {}^h\theta_{pi}) + z_h \cos({}^h\theta'_p - {}^h\theta_{pi})}{x_h \cos({}^h\theta'_p - {}^h\theta_{pi}) - z_h \sin({}^h\theta'_p - {}^h\theta_{pi}) - d_x} \right| \quad (44)$$

Using Equations (37) and (38), ${}^h\theta_{pq}$ and ${}^w\theta_{pq}$ can be obtained.

3.2.3. Vertical Rotation Angle Calculation

When the target's y coordinate in the head coordinate system is greater than d_y , the current angle of the head in the vertical direction is ${}^h\theta_{ti}$, and when the head is rotated in the vertical direction to ${}^h\theta'_t$, the target is in the new head coordinate system. The 3D coordinates are

$$\begin{pmatrix} x'_h \\ y'_h \\ z'_h \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos({}^h\theta'_t - {}^h\theta_{ti}) & \sin({}^h\theta'_t - {}^h\theta_{ti}) & 0 \\ 0 & -\sin({}^h\theta'_t - {}^h\theta_{ti}) & \cos({}^h\theta'_t - {}^h\theta_{ti}) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_h \\ y_h \\ z_h \\ 1 \end{pmatrix} \quad (45)$$

Therefore,

$$\begin{pmatrix} x'_h \\ y'_h \\ z'_h \end{pmatrix} = \begin{pmatrix} x_h \\ y_h \cos({}^h\theta'_t - {}^h\theta_{ti}) + z_h \sin({}^h\theta'_t - {}^h\theta_{ti}) \\ z_h \cos({}^h\theta'_t - {}^h\theta_{ti}) - y_h \sin({}^h\theta'_t - {}^h\theta_{ti}) \end{pmatrix} \quad (46)$$

Using Equation (29), the coordinates of the eye coordinate system after the rotation of the target can be calculated:

$$\begin{pmatrix} e x'_h \\ e y'_h \\ e z'_h \end{pmatrix} = \begin{pmatrix} d_x + x_h \\ y_h \cos({}^h\theta'_t - {}^h\theta_{ti}) + z_h \sin({}^h\theta'_t - {}^h\theta_{ti}) + d_y \\ z_h \cos({}^h\theta'_t - {}^h\theta_{ti}) - y_h \sin({}^h\theta'_t - {}^h\theta_{ti}) \end{pmatrix} \quad (47)$$

After rotation, the left and right motion modules reach the rotatable maximum value ${}^e\theta_{tmax}$ in the vertical direction, so that

$$\tan({}^e\theta_{tmax}) = \frac{e y'_h}{e z'_h} = \frac{y_h \cos({}^h\theta'_t - {}^h\theta_{ti}) + z_h \sin({}^h\theta'_t - {}^h\theta_{ti}) + d_y}{z_h \cos({}^h\theta'_t - {}^h\theta_{ti}) - y_h \sin({}^h\theta'_t - {}^h\theta_{ti})} \quad (48)$$

Simplifying Equation (48), we obtain

$$\sin({}^h\theta'_t - {}^h\theta_{ti}) = \frac{z_h \tan({}^e\theta_{tmax}) - y_h}{z_h + y_h \tan({}^e\theta_{tmax})} \cos({}^h\theta'_t - {}^h\theta_{ti}) - \frac{d_y}{z_h + y_h \tan({}^e\theta_{tmax})} \quad (49)$$

Let

$$\begin{cases} k_1 = \frac{z_h \tan({}^e\theta_{tmax}) - y_h}{z_h + y_h \tan({}^e\theta_{tmax})} \\ k_2 = -\frac{d_y}{z_h + y_h \tan({}^e\theta_{tmax})} \end{cases} \quad (50)$$

Therefore,

$${}^h\theta'_t = {}^h\theta_{ti} + \arccos\left(\frac{-k_1k_2 \pm \sqrt{k_1^2 - k_2^2 + 1}}{k_1^2 + 1}\right) \quad (51)$$

Equation (51) has two solutions; therefore, we choose the solution in which the deviation e of Equation (52) is minimized:

$$e = \left| \tan({}^e\theta_{t\max}) - \frac{y_h \cos({}^h\theta'_t - {}^h\theta_{ti}) + z_h \sin({}^h\theta'_t - {}^h\theta_{ti}) + d_y}{z_h \cos({}^h\theta'_t - {}^h\theta_{ti}) - y_h \sin({}^h\theta'_t - {}^h\theta_{ti})} \right| \quad (52)$$

Similarly, when the target's y coordinates in the head coordinate system are less than d_y , we have

$$\tan(-{}^e\theta_{t\max}) = \frac{{}^e y'_h}{{}^e z'_h} = \frac{y_h \cos({}^h\theta'_t - {}^h\theta_{ti}) + z_h \sin({}^h\theta'_t - {}^h\theta_{ti}) + d_y}{z_h \cos({}^h\theta'_t - {}^h\theta_{ti}) - y_h \sin({}^h\theta'_t - {}^h\theta_{ti})} \quad (53)$$

$$\sin({}^h\theta'_t - {}^h\theta_{ti}) = -\frac{z_h \tan({}^e\theta_{t\max}) + y_h}{z_h - y_h \tan({}^e\theta_{t\max})} \cos({}^h\theta'_t - {}^h\theta_{ti}) - \frac{d_y}{z_h - y_h \tan({}^e\theta_{t\max})} \quad (54)$$

$$\begin{cases} k_1 = -\frac{z_h \tan({}^e\theta_{t\max}) + y_h}{z_h - y_h \tan({}^e\theta_{t\max})} \\ k_2 = -\frac{d_y}{z_h - y_h \tan({}^e\theta_{t\max})} \end{cases} \quad (55)$$

$$e = \left| -\tan({}^e\theta_{t\max}) - \frac{y_h \cos({}^h\theta'_t - {}^h\theta_{ti}) + z_h \sin({}^h\theta'_t - {}^h\theta_{ti}) + d_y}{z_h \cos({}^h\theta'_t - {}^h\theta_{ti}) - y_h \sin({}^h\theta'_t - {}^h\theta_{ti})} \right| \quad (56)$$

When the obtained ${}^h\theta'_t$ is outside of the range $[-{}^h\theta_{t\max}, {}^h\theta_{t\max}]$, the value of ${}^h\theta_{tq}$ is

$${}^h\theta_{tq} = \begin{cases} {}^h\theta_{t\max} & , \quad {}^h\theta'_t \geq {}^h\theta_{t\max} \\ -{}^h\theta_{t\max} & , \quad {}^h\theta'_t \leq -{}^h\theta_{t\max} \\ {}^h\theta'_t & , \quad \text{else} \end{cases} \quad (57)$$

After obtaining ${}^h\theta_{pq}$, ${}^h\theta_{tq}$ and ${}^w\theta_{pq}$, $\mathbf{P}'_e(x'_e, y'_e, z'_e)$ are the coordinates of the target in the eye coordinate system after the mobile robot and the head are rotated:

$$\begin{pmatrix} x'_e \\ y'_e \\ z'_e \\ 1 \end{pmatrix} = \begin{pmatrix} \text{Rot}(X, {}^h\theta_{tq}) \text{Rot}(Y, {}^h\theta_{pq}) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \text{Rot}(Y, {}^w\theta_{pq}) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} + \begin{pmatrix} d_x \\ d_y \\ 0 \\ 0 \end{pmatrix} \quad (58)$$

The desired observation pose of the eye, characterized by ${}^1\theta_{tq}$, ${}^1\theta_{pq}$, ${}^r\theta_{tq}$ and ${}^r\theta_{pq}$, can be obtained using the method described in the following section.

3.2.4. Calculation of the Desired Observation Poses of the Eye

According to Formula (2), ${}^1\theta_{tq} = {}^r\theta_{tq} = \theta_t$, and ${}^1\theta_{pq} = {}^r\theta_{pq} = \theta_p$.

The inverse of the hand-eye matrix of the left camera and left motion module end coordinate system is

$${}^1T_m^{-1} = \begin{pmatrix} {}^1n_x & {}^1o_x & {}^1a_x & {}^1p_x \\ {}^1n_y & {}^1o_y & {}^1a_y & {}^1p_y \\ {}^1n_z & {}^1o_z & {}^1a_z & {}^1p_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (59)$$

The coordinate ${}^1P_c({}^1x_c, {}^1y_c, {}^1z_c)$ of $\mathbf{P}'_e(x'_e, y'_e, z'_e)$ in the left camera coordinate system satisfies the following relationship:

$$\begin{pmatrix} {}^1P_c \\ 1 \end{pmatrix} = {}^1T_m^{-1} \begin{pmatrix} \text{Rot}(X, -{}^1\theta_t) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \text{Rot}(Y, -{}^1\theta_p) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{P}'_e \\ 1 \end{pmatrix} \quad (60)$$

According to the small hole imaging model, the imaging coordinates of the $\mathbf{P}'_e(x'_e, y'_e, z'_e)$ point in the left camera are

$$\begin{pmatrix} {}^1u \\ {}^1v \\ 1 \end{pmatrix} = {}^1M_{\text{in}} {}^1P_{1c} = \begin{pmatrix} {}^1k_x & 0 & {}^1u_0 \\ 0 & {}^1k_y & {}^1v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} {}^1x_c/{}^1z_c \\ {}^1y_c/{}^1z_c \\ 1 \end{pmatrix} = \begin{pmatrix} {}^1k_x {}^1x_c/{}^1z_c + {}^1u_0 \\ {}^1k_y {}^1y_c/{}^1z_c + {}^1v_0 \\ 1 \end{pmatrix} \quad (61)$$

Substituting Equation (61) into Equation (2), we obtain

$$\begin{pmatrix} {}^1\Delta u \\ {}^1\Delta v \\ 1 \end{pmatrix} = \begin{pmatrix} {}^1k_x {}^1x_c/{}^1z_c \\ {}^1k_y {}^1y_c/{}^1z_c \\ 1 \end{pmatrix} \quad (62)$$

Based on the same principle, the coordinate ${}^rP_c(x_c, y_c, z_c)$ of $\mathbf{P}'_e(x'_e, y'_e, z'_e)$ in the right camera coordinate system is

$$\begin{pmatrix} {}^rP_c \\ 1 \end{pmatrix} = {}^rT_m^{-1} \begin{pmatrix} \text{Rot}(X, -\theta_t) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \text{Rot}(Y, -\theta_p) & 0 \\ 0 & 1 \end{pmatrix} {}^rT_e \begin{pmatrix} \mathbf{P}'_e \\ 1 \end{pmatrix} \quad (63)$$

The imaging coordinates of point $\mathbf{P}'_e(x'_e, y'_e, z'_e)$ in the right camera are

$$\begin{pmatrix} {}^ru \\ {}^rv \\ 1 \end{pmatrix} = {}^rM_{\text{in}} {}^rP_{1c} = \begin{pmatrix} {}^rk_x & 0 & {}^ru_0 \\ 0 & {}^rk_y & {}^rv_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} {}^rx_c/{}^rz_c \\ {}^ry_c/{}^rz_c \\ 1 \end{pmatrix} = \begin{pmatrix} {}^rk_x {}^rx_c/{}^rz_c + {}^ru_0 \\ {}^rk_y {}^ry_c/{}^rz_c + {}^rv_0 \\ 1 \end{pmatrix} \quad (64)$$

$$\begin{pmatrix} {}^r\Delta u \\ {}^r\Delta v \\ 1 \end{pmatrix} = \begin{pmatrix} {}^rk_x {}^rx_c/{}^rz_c \\ {}^rk_y {}^ry_c/{}^rz_c \\ 1 \end{pmatrix} \quad (65)$$

By Equations (2), (62) and (65), two equations related to θ_t and θ_p (see Appendix C for the complete equations) can be obtained. It is challenging to calculate the values of θ_t and θ_p directly from these two equations, however. To obtain a solution, we consider a suboptimal observation pose and use this pose as the initial value; then, we use the trial-and-error method to obtain the optimal observation pose. When θ_t is calculated, let $\theta_p = 0$; the solution of θ_t can then be obtained by $\Delta v_l = -\Delta v_r$. When θ_p is calculated, the solution of θ_p is solved by $\Delta u_l = -\Delta u_r$. The solution $\mathbf{P}_{\text{ob}}(\theta_t, \theta_t, \theta_p, \theta_p)$ is a suboptimal observed pose. Based on the suboptimal observation pose, the trial-and-error method can be used to obtain the optimal solution with the smallest error. The range of θ_t is $[-\theta_{t\text{max}}, \theta_{t\text{max}}]$. The range of θ_p is $[-\theta_{p\text{max}}, \theta_{p\text{max}}]$.

According to Equations (60) and (63), let θ_p be equal to 0 to obtain

$$\begin{pmatrix} {}^1P_c \\ 1 \end{pmatrix} = ({}^1T_m)^{-1} \begin{pmatrix} \text{Rot}(X, -\theta_t) & 0 \\ 0 & 1 \end{pmatrix} {}^1T_e \begin{pmatrix} \mathbf{P}'_e \\ 1 \end{pmatrix} \quad (66)$$

The following result is also available:

$$\begin{pmatrix} {}^rP_c \\ 1 \end{pmatrix} = ({}^rT_m)^{-1} \begin{pmatrix} \text{Rot}(X, -\theta_t) & 0 \\ 0 & 1 \end{pmatrix} {}^rT_e \begin{pmatrix} \mathbf{P}'_e \\ 1 \end{pmatrix} \quad (67)$$

The base coordinate system of the left motion module is the world coordinate system. Therefore, 1T_w is a unit matrix. To simplify the calculation, we have

$$\begin{pmatrix} {}^rP_e \\ 1 \end{pmatrix} = \begin{pmatrix} {}^rx'_e \\ {}^ry'_e \\ {}^rz'_e \\ 1 \end{pmatrix} = {}^rT_e \begin{pmatrix} \mathbf{P}'_e \\ 1 \end{pmatrix} \quad (68)$$

According to the calculation principle of Section 3.2.1, we have the following:

$$\begin{cases} \Delta u_l = f_x \frac{{}^l o_x y'_e + {}^l a_x z'_e}{({}^l o_z y'_e + {}^l a_z z'_e)} \cos \theta_t + \frac{{}^l o_x z'_e - {}^l a_x y'_e}{({}^l o_z z'_e - {}^l a_z y'_e)} \sin \theta_t + \frac{{}^l p_x + {}^l n_x x'_e}{({}^l p_y + {}^l n_y y'_e)} \\ \Delta v_l = f_y \frac{{}^l o_y y'_e + {}^l a_y z'_e}{({}^l o_z y'_e + {}^l a_z z'_e)} \cos \theta_t + \frac{{}^l o_y z'_e - {}^l a_y y'_e}{({}^l o_z z'_e - {}^l a_z y'_e)} \sin \theta_t + \frac{{}^l p_y + {}^l n_y y'_e}{({}^l p_z + {}^l n_z z'_e)} \end{cases} \quad (69)$$

$$\begin{cases} \Delta u_r = r f_x \frac{{}^r o_x y'_e + {}^r a_x z'_e}{({}^r o_z y'_e + {}^r a_z z'_e)} \cos \theta_t + \frac{{}^r o_x z'_e - {}^r a_x y'_e}{({}^r o_z z'_e - {}^r a_z y'_e)} \sin \theta_t + \frac{{}^r p_x + {}^r n_x x'_e}{({}^r p_y + {}^r n_y y'_e)} \\ \Delta v_r = r f_y \frac{{}^r o_y y'_e + {}^r a_y z'_e}{({}^r o_z y'_e + {}^r a_z z'_e)} \cos \theta_t + \frac{{}^r o_y z'_e - {}^r a_y y'_e}{({}^r o_z z'_e - {}^r a_z y'_e)} \sin \theta_t + \frac{{}^r p_y + {}^r n_y y'_e}{({}^r p_z + {}^r n_z z'_e)} \end{cases} \quad (70)$$

Assume the following:

$$E_{sv} = |\Delta v_l| + |\Delta v_r| \quad (71)$$

The solution to θ_t that keeps the target at the center of the two cameras needs to satisfy the following conditions:

$$\begin{cases} \Delta v_l + \Delta v_r = 0 \\ -\theta_{tmax} \leq \theta_t \leq \theta_{tmax} \\ \theta_t = \operatorname{argmin}(E_{sv}) \end{cases} \quad (72)$$

Substituting the second equation of Equations (69) and (70) into Equation (72) and solving the equation, we have

$$k_1 \cos^2 \theta_t + k_2 \sin^2 \theta_t + k_3 \sin \theta_t \cos \theta_t + k_4 \cos \theta_t + k_5 \sin \theta_t + k_6 = 0 \quad (73)$$

where k_1, k_2, k_3, k_4, k_5 are

$$k_1 = f_y ({}^l o_y y'_e + {}^l a_y z'_e) ({}^r o_z y'_e + {}^r a_z z'_e) + r f_y ({}^l o_z y'_e + {}^l a_z z'_e) ({}^r o_y y'_e + {}^r a_y z'_e) \quad (74)$$

$$k_2 = f_y ({}^l o_y z'_e - {}^l a_y y'_e) ({}^r o_z z'_e - {}^r a_z y'_e) + r f_y ({}^l o_z z'_e - {}^l a_z y'_e) ({}^r o_y z'_e - {}^r a_y y'_e) \quad (75)$$

$$k_3 = f_y ({}^l o_y y'_e + {}^l a_y z'_e) ({}^r o_z z'_e - {}^r a_z y'_e) + f_y ({}^l o_y z'_e - {}^l a_y y'_e) ({}^r o_z y'_e + {}^r a_z z'_e) + r f_y ({}^l o_z y'_e + {}^l a_z z'_e) ({}^r o_y z'_e - {}^r a_y y'_e) + r f_y ({}^l o_z z'_e - {}^l a_z y'_e) ({}^r o_y y'_e + {}^r a_y z'_e) \quad (76)$$

$$k_4 = f_y ({}^l o_y y'_e + {}^l a_y z'_e) ({}^r p_z + {}^r n_z x'_e) + f_y ({}^l p_y + {}^l n_y x'_e) ({}^r o_z y'_e + {}^r a_z z'_e) + r f_y ({}^l o_z y'_e + {}^l a_z z'_e) ({}^r p_y + {}^r n_y x'_e) + r f_y ({}^l p_z + {}^l n_z x'_e) ({}^r o_y y'_e + {}^r a_y z'_e) \quad (77)$$

$$k_5 = f_y ({}^l o_y z'_e - {}^l a_y y'_e) ({}^r p_z + {}^r n_z x'_e) + f_y ({}^l p_y + {}^l n_y x'_e) ({}^r o_z z'_e - {}^r a_z y'_e) + r f_y ({}^l o_z z'_e - {}^l a_z y'_e) ({}^r p_y + {}^r n_y x'_e) + r f_y ({}^l p_z + {}^l n_z x'_e) ({}^r o_y z'_e - {}^r a_y y'_e) \quad (78)$$

$$k_6 = f_y ({}^l p_y + {}^l n_y x'_e) ({}^r p_z + {}^r n_z x'_e) + r f_y ({}^l p_z + {}^l n_z x'_e) ({}^r p_y + {}^r n_y x'_e) \quad (79)$$

According to the triangle relationship, we have

$$\cos^2 \theta_t + \sin^2 \theta_t = 1 \quad (80)$$

Replacing $\cos \theta_t$ in Equation (73) with $\sin \theta_t$, we obtain the following:

$$k'_1 \sin^4 \theta_t + k'_2 \sin^3 \theta_t + k'_3 \sin^2 \theta_t + k'_4 \sin \theta_t + k'_5 = 0 \quad (81)$$

where k_1, k_2, k_3, k_4, k_5 are

$$k'_1 = (k_2 - k_1)^2 + k_3^2 \quad (82)$$

$$k'_2 = 2(k_2 - k_1)k_5 + 2k_3k_4 \quad (83)$$

$$k'_3 = 2(k_2 - k_1)k_6 + k_5^2 + k_4^2 - k_3^2 \quad (84)$$

$$k'_4 = 2k_5k_6 - 2k_3k_4 \quad (85)$$

$$k'_5 = k_6^2 - k_4^2 \quad (86)$$

Four solutions can be obtained using Equation (81). The optimal solution is a real number, and the most suitable solution can be selected by the condition of Equation (72).

After θ_t is obtained, θ_p can be solved based on the obtained θ_t .

According to Equations (60) and (63), $\bar{\theta}_t$ is the solution obtained in Section 3.2.2, so that

$$\begin{pmatrix} {}^1P_c \\ 1 \end{pmatrix} = ({}^1T_m)^{-1} \begin{pmatrix} \text{Rot}(X, -\bar{\theta}_t) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \text{Rot}(Y, -{}^1\theta_p) & 0 \\ 0 & 1 \end{pmatrix} {}^1T_e \begin{pmatrix} P'_e \\ 1 \end{pmatrix} \quad (87)$$

The following result is also available:

$$\begin{pmatrix} {}^rP_c \\ 1 \end{pmatrix} = ({}^rT_m)^{-1} \begin{pmatrix} \text{Rot}(X, -\bar{\theta}_t) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \text{Rot}(Y, -{}^r\theta_p) & 0 \\ 0 & 1 \end{pmatrix} {}^rT_e \begin{pmatrix} P'_e \\ 1 \end{pmatrix} \quad (88)$$

Since $\bar{\theta}_t$ is known, for convenience of calculation, we set

$${}^1T'_m = ({}^1T_m)^{-1} \begin{pmatrix} \text{Rot}(X, -\bar{\theta}_t) & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} {}^1n'_x & {}^1o'_x & {}^1a'_x & {}^1p'_x \\ {}^1n'_y & {}^1o'_y & {}^1a'_y & {}^1p'_y \\ {}^1n'_z & {}^1o'_z & {}^1a'_z & {}^1p'_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (89)$$

$${}^rT'_m = ({}^rT_m)^{-1} \begin{pmatrix} \text{Rot}(X, -\bar{\theta}_t) & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} {}^rn'_x & {}^ro'_x & {}^ra'_x & {}^rp'_x \\ {}^rn'_y & {}^ro'_y & {}^ra'_y & {}^rp'_y \\ {}^rn'_z & {}^ro'_z & {}^ra'_z & {}^rp'_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (90)$$

The following results are obtained:

$$\begin{cases} \Delta u_l = {}^1f_x \frac{({}^1n'_x x'_e + {}^1a'_x z'_e) \cos \theta_p + ({}^1a'_x x'_e - {}^1n'_x z'_e) \sin \theta_p + ({}^1p'_x + {}^1o'_x y'_e)}{({}^1n'_z x'_e + {}^1a'_z z'_e) \cos \theta_p + ({}^1a'_z x'_e - {}^1n'_z z'_e) \sin \theta_p + ({}^1p'_z + {}^1o'_z y'_e)} \\ \Delta v_l = {}^1f_y \frac{({}^1n'_y x'_e + {}^1a'_y z'_e) \cos \theta_p + ({}^1a'_y x'_e - {}^1n'_y z'_e) \sin \theta_p + ({}^1p'_y + {}^1o'_y y'_e)}{({}^1n'_z x'_e + {}^1a'_z z'_e) \cos \theta_p + ({}^1a'_z x'_e - {}^1n'_z z'_e) \sin \theta_p + ({}^1p'_z + {}^1o'_z y'_e)} \end{cases} \quad (91)$$

$$\begin{cases} \Delta u_r = {}^rf_x \frac{({}^rn'_x x'_e + {}^ra'_x z'_e) \cos \theta_p + ({}^ra'_x x'_e - {}^rn'_x z'_e) \sin \theta_p + ({}^rp'_x + {}^ro'_x y'_e)}{({}^rn'_z x'_e + {}^ra'_z z'_e) \cos \theta_p + ({}^ra'_z x'_e - {}^rn'_z z'_e) \sin \theta_p + ({}^rp'_z + {}^ro'_z y'_e)} \\ \Delta v_r = {}^rf_y \frac{({}^rn'_y x'_e + {}^ra'_y z'_e) \cos \theta_p + ({}^ra'_y x'_e - {}^rn'_y z'_e) \sin \theta_p + ({}^rp'_y + {}^ro'_y y'_e)}{({}^rn'_z x'_e + {}^ra'_z z'_e) \cos \theta_p + ({}^ra'_z x'_e - {}^rn'_z z'_e) \sin \theta_p + ({}^rp'_z + {}^ro'_z y'_e)} \end{cases} \quad (92)$$

Assume that

$$E_{su} = |\Delta u_l| + |\Delta u_r| \quad (93)$$

The solution to θ_p that keeps the target at the center of the two cameras needs to satisfy the following conditions:

$$\begin{cases} \Delta u_l + \Delta u_r = 0 \\ -\theta_{pmax} \leq \theta_p \leq \theta_{pmax} \\ \theta_p = \text{argmin}(E_{su}) \end{cases} \quad (94)$$

Substituting the second equation of Equations (91) and (92) into Equation (94) and solving the available equation, we obtain

$$k_1 \cos^2 \theta_p + k_2 \sin^2 \theta_p + k_3 \sin \theta_p \cos \theta_p + k_4 \cos \theta_p + k_5 \sin \theta_p + k_6 = 0 \quad (95)$$

where

$$k_1 = {}^1f_x ({}^1n'_x x'_e + {}^1a'_x z'_e) ({}^rn'_z x'_e + {}^ra'_z z'_e) + {}^rf_x ({}^rn'_x x'_e + {}^ra'_x z'_e) ({}^1n'_z x'_e + {}^1a'_z z'_e) \quad (96)$$

$$k_2 = {}^1f_x({}^1n'_x z'_e - {}^1a'_x x'_e)({}^r n'_z r z'_e - {}^r a'_z r x'_e) + {}^r f_x({}^r n'_x r z'_e - {}^r a'_x r x'_e)({}^1 n'_z z'_e - {}^1 a'_z x'_e) \quad (97)$$

$$k_3 = {}^1f_x({}^1n'_x x'_e + {}^1a'_x z'_e)({}^r n'_z r z'_e - {}^r a'_z r x'_e) + {}^1f_x({}^1n'_x z'_e - {}^1a'_x x'_e)({}^r n'_z r x'_e + {}^r a'_z r z'_e) \\ + {}^r f_x({}^r n'_x r z'_e - {}^r a'_x r x'_e)({}^1 n'_z x'_e + {}^1 a'_z z'_e) + {}^r f_x({}^r n'_x r x'_e + {}^r a'_x r z'_e)({}^1 n'_z z'_e - {}^1 a'_z x'_e) \quad (98)$$

$$k_4 = {}^1f_x({}^1n'_x x'_e + {}^1a'_x z'_e)({}^r o'_z r y'_e + {}^r p'_z) + {}^1f_x({}^1o'_x y'_e + {}^1p'_x)({}^r n'_z r x'_e + {}^r a'_z r z'_e) \\ + {}^r f_x({}^r o'_x r y'_e + {}^r p'_x)({}^1 n'_z x'_e + {}^1 a'_z z'_e) + {}^r f_x({}^r n'_x r x'_e + {}^r a'_x r z'_e)({}^1 o'_z y'_e + {}^1 p'_z) \quad (99)$$

$$k_5 = {}^1f_x({}^1n'_x z'_e - {}^1a'_x x'_e)({}^r o'_z r y'_e + {}^r p'_z) + {}^1f_x({}^1o'_x y'_e + {}^1p'_x)({}^r n'_z r z'_e - {}^r a'_z r x'_e) \\ + {}^r f_x({}^r o'_x r y'_e + {}^r p'_x)({}^1 n'_z z'_e - {}^1 a'_z x'_e) + {}^r f_x({}^r n'_x r z'_e - {}^r a'_x r x'_e)({}^1 o'_z y'_e + {}^1 p'_z) \quad (100)$$

$$k_6 = {}^1f_x({}^1o'_x y'_e + {}^1p'_x)({}^r o'_z r y'_e + {}^r p'_z) + {}^r f_x({}^r o'_x r y'_e + {}^r p'_x)({}^1 o'_z y'_e + {}^1 p'_z) \quad (101)$$

Replacing $\cos\theta_p$ in Equation (73) with $\sin\theta_p$, we obtain

$$k'_1 \sin^4 \theta_p + k'_2 \sin^3 \theta_p + k'_3 \sin^2 \theta_p + k'_4 \sin \theta_p + k'_5 = 0 \quad (102)$$

where

$$k'_1 = (k_2 - k_1)^2 + (k_3)^2 \quad (103)$$

$$k'_2 = 2(k_2 - k_1)k_5 + 2k_3k_4 \quad (104)$$

$$k'_3 = 2(k_2 - k_1)k_6 + (k_5)^2 + (k_4)^2 - (k_3)^2 \quad (105)$$

$$k'_4 = 2k_5k_6 - 2k_3k_4 \quad (106)$$

$$k'_5 = (k_6)^2 - (k_4)^2 \quad (107)$$

Four solutions can be obtained using Equation (102). The optimal solution must be a real number, and the most suitable solution can be selected using the condition of Equation (94). For the case where the four solutions cannot satisfy Equation (94), the position of the target is beyond the position that the bionic eye can reach. In this case, compensation is required through the head or torso. $\bar{\theta}_t$ and $\bar{\theta}_p$ obtained at this time are suboptimal solutions close to the optimal solution. θ_t and θ_p are the optimal solutions.

Through the above steps, the desired observation pose can be calculated. The calculation steps of ${}^s f_q$ can be summarized by the flow chart shown in Figure 8.

3.3. Desired Pose Calculation for Approaching Gaze Point Tracking

The mobile robot approaches the target in two steps: the first step is that the robot and the head rotate in the horizontal direction until the robot and the head are facing the target, and the second step is that the robot moves straight toward the target. The desired position of the approaching motion should satisfy the following conditions: (1) the target should be on the Z axis of the robot and the head coordinate system, (2) the distance between the target and the robot should be less than the set threshold D_T and (3) the eye should be in the optimal observation position. ${}^a f_q$ can be defined as

$${}^a f_q = \begin{pmatrix} {}^w x_{mq} = 0 \\ {}^w z_{mq} = \{z \mid 0 < {}^w z_m - z \leq D_T\} \\ {}^w \theta_{pq} = \{\theta \mid |\theta| \leq 2\pi, {}^w \gamma_p = 0\} \\ {}^h \theta_{pq} = 0 \\ {}^h \theta_{tq} = \{\theta \mid |\theta| \leq {}^h \theta_{tmax}, |{}^h \gamma_t| \leq {}^h \gamma_{tmax}\} \\ {}^l \theta_{pq} = {}^r \theta_{pq} = \{\theta \mid |\theta| \leq {}^e \theta_{pmax}, \Delta m_1 = -\Delta m_r\} \\ {}^l \theta_{tq} = {}^r \theta_{tq} = \{\theta \mid |\theta| \leq {}^e \theta_{tmax}, \Delta m_1 = -\Delta m_r\} \end{pmatrix} \quad (108)$$

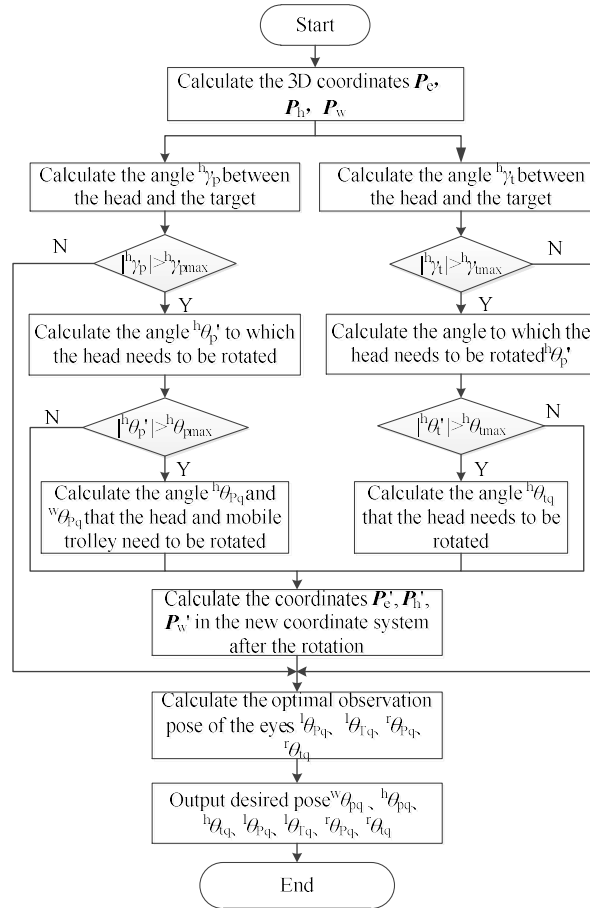


Figure 8. Steps for calculating the desired pose of the fixation point.

The desired rotation angle ${}^w \theta_{pq}$ of the moving robot is the same as the angle ${}^b \gamma_p$ between the robot and the target and can be obtained by

$${}^w \theta_{pq} = {}^w \gamma_p = \arctan\left(\frac{{}^w z_m}{{}^w x_m}\right) \quad (109)$$

${}^h \theta_{tq}$ can be obtained using the method described in Section 3.2. The optimal observation pose described in Section 3.2.4 can be used to obtain ${}^l \theta_{tq}$, ${}^l \theta_{pq}$, ${}^r \theta_{tq}$ and ${}^r \theta_{pq}$.

4. Robot Pose Control

After obtaining the desired pose of the robot system, the control block diagram shown in Figure 9 is used to control the robot to move to the desired pose.

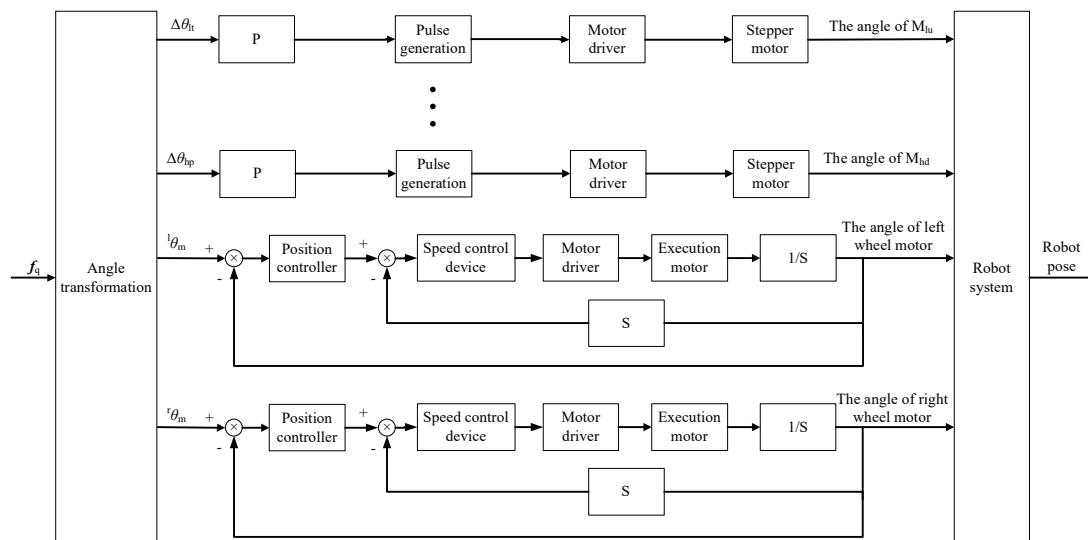


Figure 9. Robot pose control block diagram.

The desired pose is converted to the desired position of the motor. $\Delta\theta_{lt}$, $\Delta\theta_{lp}$, $\Delta\theta_{rt}$, $\Delta\theta_{rp}$, $\Delta\theta_{lt}$ and $\Delta\theta_{hp}$ are deviations of the desired angle from the current angle of motor M_{lu} , motor M_{ld} , motor M_{ru} , motor M_{rd} , motor M_{hu} and motor M_{hd} , respectively. ${}^l\theta_m$ and ${}^r\theta_m$ are the angles at which each wheel of the moving robot needs to be rotated. During the in situ gaze point tracking process, the moving robot performs only the rotation of the original position, and the angle of the robot movement can be calculated according to the desired angle of the robot. When the robot rotates, the two wheels move in opposite directions at the same speed. Let the distance between the two wheels of the moving robot be D_r ; when the robot rotates around an angle ${}^w\theta_{pq}$, the distance that each wheel needs to move is

$$S = {}^w\theta_{pq} \frac{D_r}{2} \quad (110)$$

The diameter of each wheel is d_w , and the angle of rotation of each wheel is (where counterclockwise is positive)

$${}^r\theta_m = -{}^l\theta_m = \frac{2S}{d_w} \quad (111)$$

In the process of approaching the target, the moving robot follows a straight line, and the angle of rotation of each wheel is

$${}^r\theta_m = {}^l\theta_m = \frac{2{}^wz_{mq}}{d_w} \quad (112)$$

The movement of the moving robot is achieved by controlling the rotation of each wheel. Each wheel is equipped with a DC brushless motor, and a DSP2000 controller is used to control the movement of the DC brushless motor. Position servo control is implemented in the DSP2000 controller.

In the robot system, the weight of the camera and lens is approximately 80 g, the weight of the camera and the fixed mechanical parts is approximately 50 g and the motor that controls the vertical rotation of the camera (rotating around the horizontal axis of rotation) and the corresponding encoder weighs approximately 250 g. The mechanical parts of the fixed vertical rotating motor and encoder weigh approximately 100 g. The radius of the rotation of the camera in the vertical direction is approximately 1 cm, and the rotation in the horizontal direction (rotation about the vertical axis of rotation) has a radius of approximately 2 cm. Therefore, when the gravitational acceleration is 9.8 m/s^2 , the torque required for the vertical rotating electric machine is approximately 0.013 N·m, and the torque required for the horizontal rotating electric machine is approximately 0.043 N·m. The

vertical rotating motor uses a 28BYG5401 stepping motor with a holding torque of 0.1 N·m and a positioning torque of 0.008 N·m. The driver is HSM20403A. The horizontal rotating motor is a 57BYGH301 stepping motor with a holding torque of 1.5 N·m, a positioning torque of 0.07 N·m and drive model HSM20504A. The four stepping motors of the eye have a step angle of 1.8° and are all subdivided by 25, so the actual step angle of each motor is 0.072°, and the minimum pulse width that the driver can receive is 2.5 μs. The stepper motor has a maximum angular velocity of 200°/s.

The head vertical rotary motor uses a 57BYGH401 stepper motor with a holding torque of 2.2 N·m, a positioning torque of 0.098 N·m and drive model HSM20504A. The head horizontal rotary motor is an 86BYG350B three-phase AC stepping motor with a holding torque of 5 N·m, a positioning torque of 0.3 N·m and an HSM30860M driver. The step angle of the head motor after subdivision is also 0.072°. The head vertical motor has a load of approximately 5 kg and a radius of rotation of less than 1 cm. The head horizontal rotary motor has a load of approximately 9.5 kg and a radius of rotation of approximately 5 cm. In the experiment, we found that the maximum horizontal pulse frequency that the head horizontal rotary motor can receive is 0.6 Kpps. Its maximum angular velocity is 43.2°/s.

5. Experiments and Discussion

Using the robot platform introduced in Section 2, experiments on in situ gaze point tracking and approaching gaze point tracking were performed

Each camera has a resolution of 400 × 300 pixels. The directions of rotation are [−45°, 45°]. The range of rotation of the head is [−30°, 30°]. d_x and d_y are 150 mm and 200 mm, respectively. The internal and external parameters, distortion parameters, initial position parameters and left- and right-hand-eye parameters of the dual purpose method are calibrated as follows:

$${}^lM_{in} = \begin{pmatrix} 341.58 & 0 & 201.6 \\ 0 & 341.97 & 147.62 \\ 0 & 0 & 1 \end{pmatrix} \quad (113)$$

$$K_l = (-0.1905 \quad 0.2171 \quad -0.0018 \quad -0.0005 \quad -0.0823) \quad (114)$$

$${}^lT_m = \begin{pmatrix} 1.0 & 0.0078 & -0.0022 & 58.4172 \\ 0.0001 & 0.9954 & 0.0959 & 3.6042 \\ 0.0013 & -0.0959 & 0.9954 & 51.9366 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (115)$$

$${}^rM_{in} = \begin{pmatrix} 335.13 & 0 & 184.32 \\ 0 & 335.5 & 141.26 \\ 0 & 0 & 1 \end{pmatrix} \quad (116)$$

$$K_r = (-0.1861 \quad 0.1987 \quad -0.004 \quad -0.0011 \quad -0.0739) \quad (117)$$

$${}^rT_m = \begin{pmatrix} 0.9999 & -0.0086 & -0.0125 & -45.0147 \\ 0.0190 & -0.9969 & -0.0782 & -24.5528 \\ 0.0097 & -0.0784 & 0.9970 & 42.9270 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (118)$$

$${}^lT_r = \begin{pmatrix} 0.9998 & -0.0099 & -0.0193 & 189.5922 \\ 0.0095 & 0.9997 & -0.0215 & -0.0426 \\ 0.0195 & 0.0213 & 0.9996 & 8.9671 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (119)$$

The experimental in situ gaze point tracking scene is shown in Figure 10, with a checkerboard target used as the target. For in situ gaze point tracking, the target is held by

a person. In the approaching target gaze tracking experiment, the target is fixed in front of the robot.

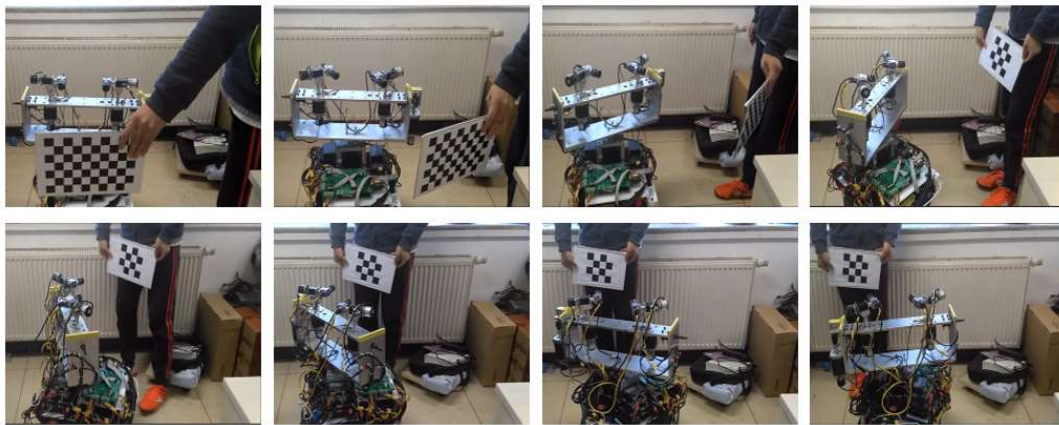


Figure 10. Experimental in situ gaze point tracking scene.

5.1. In Situ Gaze Point Tracking Experiment

In the in situ gaze experiment, the target moves at a low speed within a certain range, and the robot combines the movement of the eye, the head and the mobile robot so that the binocular vision can always perceive the 3D coordinates of the target at the optimal observation posture. This experiment prompts the robot to find the target and gaze at it. In the gaze point tracking process, binocular stereo vision is used to calculate the 3D coordinates of the target in the eye coordinate system in real time. Through the positional relationship between the eye and the head, the coordinate system of the target in the eye can be converted to the head coordinate system. Similarly, the 3D coordinates of the target in the robot coordinate system can be obtained. Through the 3D coordinates, the desired poses of the eyes, head and mobile robot are calculated according to the method proposed in this paper. Then, the camera is controlled to the desired position by the stepping motor; after reaching the desired position, the image and the motor position information are collected again, and the 3D coordinates of the target are calculated.

In the experiment, the angles between the head and the target, ${}^h\gamma_{pmax}$ and ${}^h\gamma_{pmax}$, are each 30° . The method described in Section 3 is used to calculate the desired pose of each joint of the robot based on the 3D coordinates of the target. In the experiment, the actual coordinate position and desired coordinate position of the target in the binocular image space, the actual position and desired position of the eye and head motor, the angle between the head and the robot and the target, and the target in the robot coordinate system are stored. Figure 11a,b show the u and v coordinates of the target on the left image, respectively, and Figure 11c,d show the u and v coordinates of the target on the right image, respectively. The desired image coordinates are recalculated based on the optimal observation position. Figure 11e–h show the positions of the tilt motor (M_{lu}) of the left eye, the pan motor (M_{ld}) of the left eye, the tilt motor (M_{ru}) of the right eye and the pan motor (M_{rd}) of the right eye, respectively. Figure 11i shows the positions of the pan motor (M_{hd}) of the head. Since the target moves in the vertical direction with small amplitude, the motor M_{hu} does not rotate, and the case is similar to the motion principle of the motor M_{hd} , so the motor position of the head only provides the result of the motor M_{hd} . Figure 11j shows the angle deviation and rotation. In this figure, T-h is the angle between the head and target, T-r is the angle between the robot and target, R-r is the angle of the robot rotation from the origin location and T-o is the angle of the target to the origin location. Figure 11k shows the coordinates $({}^w x, {}^w z)$ of the target in the world coordinate system. Figure 11l shows the coordinates $({}^o x, {}^o z)$ of the target in the world coordinate system of the origin location.

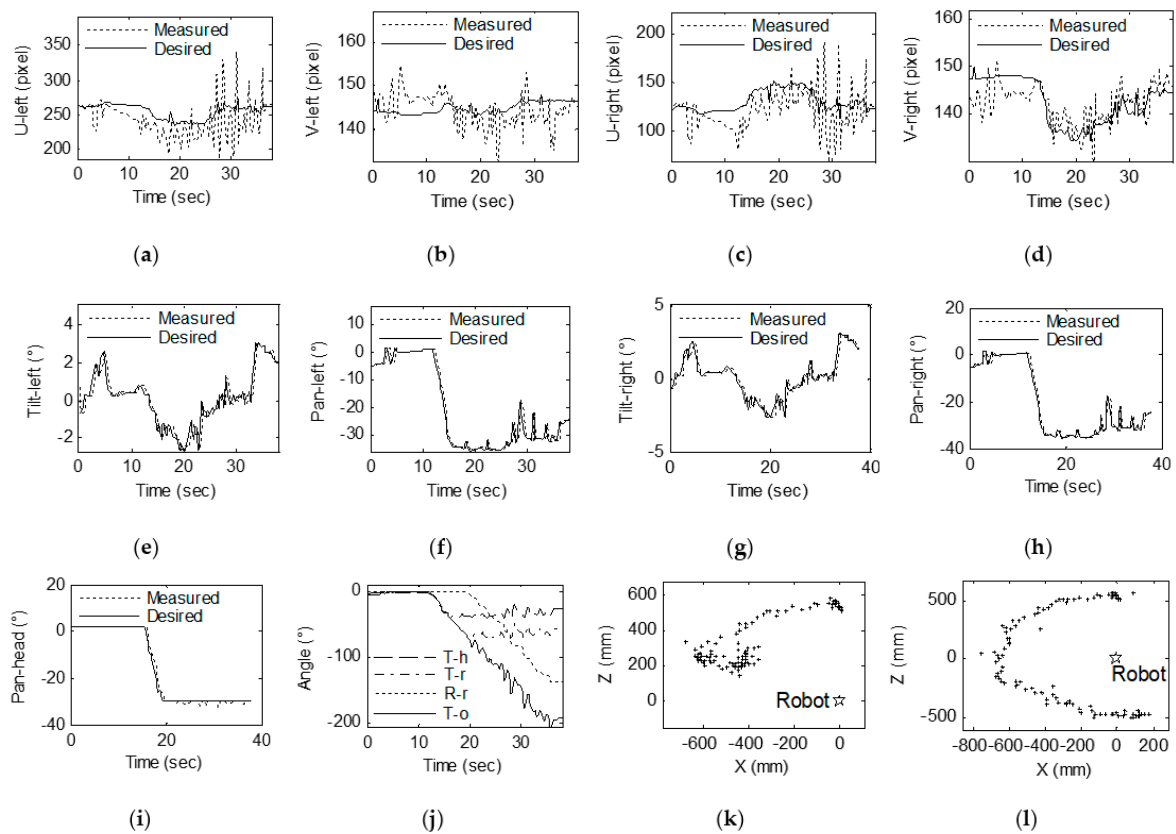


Figure 11. Experimental results of gaze shifting to the target: (a) U coordinates of the target on the left image. (b) V coordinates of the target on the left image. (c) U coordinates of the target on the right image. (d) V coordinates of the target on the right image. (e) Left camera tilt. (f) Left camera pan. (g) Right camera tilt. (h) Right camera pan. (i) Head pan. (j) Angle deviation and rotation. (k) Coordinates (${}^w x, {}^w z$) of the target in the world coordinate system. (l) Coordinates (${}^o x, {}^o z$) of the target in the world coordinate system based on the origin location. The “+” in the subfigures (k,l) represents the position of the target in the coordinate system, and the “☆” represents the position of the robot in the coordinate system.

As shown in Figure 11, the image coordinate of the target is substantially within ± 40 pixels in the central region of the left and right images in the x direction. These coordinates are kept within ± 10 pixels of the center region of the left and right images in the y direction. Throughout the experiment, the target was rotated approximately 200° around the robot. The robot moved approximately 140° , the head rotated 30° and the target could be kept in the center region of the binocular images. The motor position curve shows that the motor’s operating position can track the desired position very well. The angle variation curve shows that the angle between the target and the head and the robot changes and that the robot turning angles are suitably consistent. The coordinates of the target shown in Figure 11 in the robot coordinate system and the coordinates of the target in the initial position of the world coordinate system are very close to the actual position change in the target’s position.

Through the above analysis, we can determine the following: (1) It is feasible to realize gaze point tracking of a robot based on 3D coordinates. (2) Using the movement of the head, eyes and mobile robot used in this paper, it is possible to achieve gaze point tracking of the target while ensuring minimum resource consumption.

5.2. Approaching Gaze Point Tracking Experiment

The approaching gaze point tracking experimental scene is shown in Figure 12.

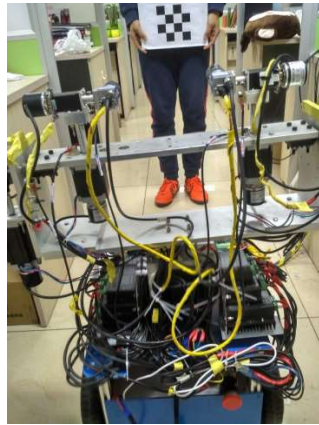


Figure 12. Experimental approaching gaze point tracking scene.

The robot approaches the target without obstacles and reaches the area in which the robot can operate on the target. The target can be grasped or carefully observed. In the approaching gaze experiment, a target is fixed at a position 2.2 m from the robot, and when the robot moves to a position where the distance from the target to robot is 0.6 m, the motion is stopped, and the maximum speed of the moving robot is 1 m/s. The experiment realizes the approaching movement to the target in two steps: first, the head, the eye and the moving robot chassis are rotated so that the head and the moving robot are facing the target, and the head observes the target in the optimal observation posture; second, the movement is controlled. The robot moves linearly in the target's direction. During the movement, the angles of the head and the eye are fine-tuned, and the 3D coordinates of the target are detected in real time until the z coordinate of the target in the robot coordinate system is less than the threshold set to stop the motion.

Figure 13 shows the results of the approaching gaze point tracking experiment. Figure 13a,b show the u and v coordinates of the target on the left image, respectively, and Figure 13c,d show the u and v coordinates of the target on the right image, respectively. The desired image coordinates are recalculated based on the optimal observation position. Figure 13e–h show the positions of the tilt motor (M_{lu}) of the left eye, the pan motor (M_{ld}) of the left eye, the tilt motor (M_{ru}) of the right eye and the pan motor (M_{rd}) of the right eye, respectively. Figure 13i shows the positions of the pan motor (M_{hd}) of the head. Figure 13j shows the angle deviation and rotation. T-h is the angle between the head and the target, T-r is the angle between the robot and the target, R-r is the angle of the robot's rotation from the origin location and T-o is the angle of the target to the origin location. Figure 13k shows the coordinates (${}^w x, {}^w z$) of the target in the world coordinate system. Figure 13l shows the robot's forward distance and the distance between the target and the robot.

The change in the image's coordinate curve indicates that the coordinates of the target in the left and right images move from the initial position to the central region of the image and stabilize in the center region of the image during the approach process. In the process of turning towards the target in the first step, the target coordinates in the image fluctuate because the head motor rotates a large amount and is accompanied by a certain vibration during the rotation, which can be avoided by using a system with better stability. The variety curve of the motor position in Figure 13 shows that the motion of the motor can track the target well with the desired pose, and the prediction of the 3D coordinates is not used during the tracking process, so this prediction is accompanied by a cycle lag. The changes in angle in Figure 13 show that the robot system achieves the task of steering toward the target in the first few control cycles and then moves toward the target at a stable angle. Figure 13a shows the change in the coordinates of the target in the robot coordinate system. When the robot rotates, fluctuations arise around the measured x coordinate, mainly due to the measurement error caused by the shaking of the system. The experimental results in Figure 13b show that the robot's movement toward the target is very consistent. During the approach process, the target can be kept within ± 50 pixels of the desired position in

the horizontal direction of the image while being within ± 20 pixels of the desired position in the vertical direction of the image. The eye motor achieves fast tracking of the target in 1.5 s. The angle between the target and the head is reduced from 20° to 0° , and the angle between the target and the robot is reduced from 35° to 0° . The robot then over-turns. At 34° , the target changes by 34° from the initial position.

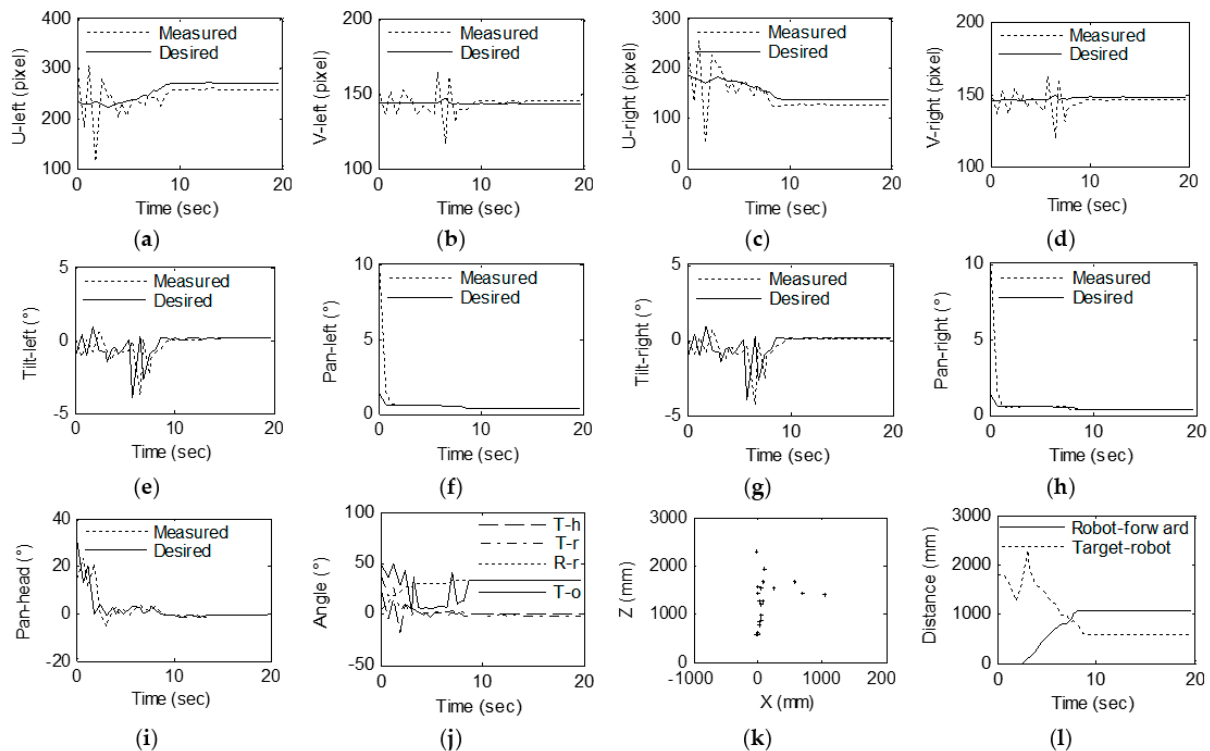


Figure 13. Experimental results of gaze shifting to the target: (a) U coordinates of the target on the left image. (b) V coordinates of the target on the left image. (c) U coordinates of the target on the right image. (d) V coordinates of the target on the right image. (e) Left camera tilt. (f) Left camera pan. (g) Right camera tilt. (h) Right camera pan. (i) Head pan. (j) Angular deviation and rotation. (k) Coordinates (${}^w x, {}^w z$) of the target in the world coordinate system. (l) Robot forward distance and the distance between the target and robot.

Through the above analysis, it can be found that by using the combination of the head, the eye and the trunk in the present method, the approach toward the target can be achieved while ensuring that the robot is gazing at the target.

6. Conclusions

This study achieved gaze point tracking based on the 3D coordinates of the target. First, a robot experiment platform was designed. Based on the bionic eye experiment platform, a head with two degrees of freedom was added, using the mobile robot as a carrier.

Based on the characteristics of the robot platform, this paper proposed a method of gaze point tracking. To achieve in situ gaze point tracking, the combination of the eyes, head and trunk is designed based on the principles of minimum resource consumption and maximum system stability. Eye rotation consumes the least amount of resources and has minimal impact on the stability of the overall system during the exercise. The head rotation consumes more resources than the eyeball but fewer than the trunk rotation. At the same time, the rotation of the head affects the stability of the eyeball but only minimally affects the stability of the entire robotic system. The resources consumed by the rotation of the trunk generally predominate, and the rotation of the trunk tends to affect the stability of the head and the eye. Therefore, when the eye can observe the target in the optimal

observation posture, only the eye is rotated; otherwise, the head is rotated, and when the angle at which the head needs to move exceeds its threshold, the mobile robot rotates. When approaching gaze point tracking is performed, the robot and head first face the target and then move straight toward the vicinity of the target. Based on the proposed gaze point tracking method, this paper provides an expected pose calculation method for the horizontal rotation angle and the vertical rotation angle.

Based on the experimental robot platform, a series of experiments was performed, and the effectiveness of the gaze point tracking method was verified. In our future works, a practical task of delivering medicine in a hospital and more detailed comparative experiments, as well as discussions with other similar studies, will be implemented.

Author Contributions: Conceptualization, Q.W. and M.Q.; Methodology, X.F.; Software, X.F.; Validation, X.F.; Formal analysis, H.C.; Investigation, M.Q.; Data curation, Q.W. and Y.Z.; Writing—original draft, X.F. and Q.W.; Writing—review & editing, M.Q.; Project administration, H.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data was created.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Measurement Error Analysis of Binocular Stereo Vision

It is assumed that the angle at which the left motion module rotates in the horizontal direction with respect to the initial position is ${}^l\theta_p$ and that the angle at which the right motion module rotates in the horizontal direction with respect to the initial position is ${}^r\theta_p$. When using the bionic eye platform for 3D coordinate calculation, we find that when ${}^l\theta_p = {}^r\theta_p$, the measured three-dimensional coordinate ratio is higher than ${}^l\theta_p > 0$ and ${}^r\theta_p < 0$, which is more accurate, as shown in Figure A1a,b. In this case, the measurement accuracy is higher than that in the case shown in Figure A1c. This chapter analyzes the measurement error of three-dimensional coordinates, explains the reason for this phenomenon and proposes a method to improve the accuracy of three-dimensional coordinate measurement.

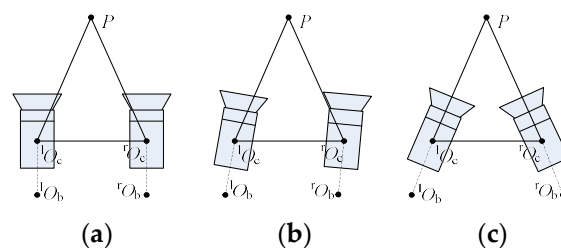


Figure A1. Binocular motion mode, (a) initial position ${}^l\theta_p = {}^r\theta_p$, (b) ${}^l\theta_p = {}^r\theta_p$, (c) ${}^l\theta_p > 0$ and ${}^r\theta_p < 0$.

For the convenience of calculation, according to the characteristics of the bionic eye platform, the binocular stereo vision measurement model shown in Figure A2 is used to analyze the error of the three-dimensional coordinate measurement.

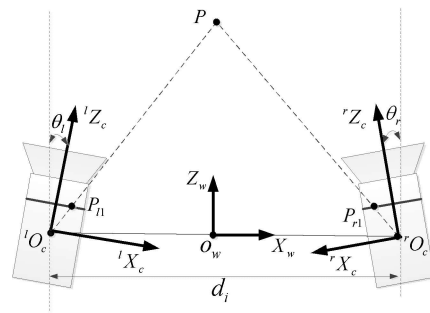


Figure A2. The principle of the vision system of bionic eyes.

Appendix A.1. Vision System of Bionic Eyes

Two cameras are used to imitate human eyes and the principle of the vision system of bionic eyes is shown in Figure 1. We suppose that the optical axes of two cameras used to imitate human eyes are coplanar. As shown in Figure 1, $O_w X_w Z_w$ is the world coordinate system, the X_w axis is along the baseline of two cameras and the Z_w axis is in the plane Π , which consists of the two cameras' optical axes. ${}^l O_c {}^l X_c {}^l Z_c$ is the coordinate system of the left camera, where ${}^l Z_c$ is along the optical axis of the left camera and ${}^l X_c$ axis is in the plane Π . ${}^r O_c {}^r X_c {}^r Z_c$ is the coordinate system of the right camera, ${}^r Z_c$ is along the optical axis of the right camera and ${}^r X_c$ axis is in the plane Π . The two cameras can move cooperatively to imitate the movement of human eyes.

Appendix A.2. Depth Measurement Model

In Figure A2, the position vector of point ${}^l O_c$ in $O_w X_w Z_w$ is ${}^w O_l = [-d_i/2, 0]^T$ and the position vector of point ${}^r O_c$ in $O_w X_w Z_w$ is ${}^w O_r = [d_i/2, 0]^T$. d_i is the length of the baseline. Let ${}^l P = [x_l, z_l]^T$ and ${}^r P = [x_r, z_r]^T$ be the position vectors of the object point P in the left and right camera coordinate systems, respectively. Let ${}^w P = [x_w, z_w]^T$ be the position vector of point P in $O_w X_w Z_w$; then, ${}^l P$ can be obtained as follows:

$${}^l P = {}^l R_w ({}^w P - {}^w O_l) \quad (A1)$$

where ${}^l R_w$ is the rotation transformation matrix from the world coordinate system to the left camera coordinate system. ${}^l R_w$ can be expressed as

$${}^l R_w = \begin{bmatrix} \cos \theta_l & \sin \theta_l \\ -\sin \theta_l & \cos \theta_l \end{bmatrix} \quad (A2)$$

where θ_l is the rotation angle which is defined as the angle between camera optical axis and the Z_w axis.

Based on the same principle, we can obtain

$${}^r P = {}^r R_w ({}^w P - {}^w O_r) \quad (A3)$$

where

$${}^r R_w = \begin{bmatrix} \cos \theta_r & \sin \theta_r \\ -\sin \theta_r & \cos \theta_r \end{bmatrix} \quad (A4)$$

As shown in Figure 1, P_{l1} is the intersection point of the line ${}^l O_c P$ and the normalized image plane of the left camera, and P_{r1} is the intersection point of the line ${}^r O_c P$ and the normalized image plane of the right camera. From the geometric relationship as shown in Figure 1, the position vector of point P_{l1} in ${}^l O_c {}^l X_c {}^l Z_c$ can be expressed as ${}^l P_{l1} = [x_l/z_l, 1]^T$, and the position vector of point P_{r1} in ${}^r O_c {}^r X_c {}^r Z_c$ can be expressed as ${}^r P_{r1} = [x_r/z_r, 1]^T$. Let ${}^w P_{l1} = [{}^w x_{l1}, {}^w z_{l1}]^T$ be the position vector of point P_{l1} in $O_w X_w Z_w$ and ${}^w P_{r1} = [{}^w x_{r1}, {}^w z_{r1}]^T$

be the position vector of point P_{r1} in $O_w X_w Z_w$. ${}^w P_{l1}$ and ${}^w P_{r1}$ can be calculated in (A5) and (A6) by coordinate transformation according to (A1) and (A3).

$${}^w P_{l1} = \left({}^l R_w\right)^{-1} {}^l P_{l1} + {}^w O_l \quad (A5)$$

$${}^w P_{r1} = \left({}^r R_w\right)^{-1} {}^r P_{r1} + {}^w O_r \quad (A6)$$

From Equations (99) and (100), we can obtain

$${}^w P_{l1} = \begin{bmatrix} {}^w x_{l1} \\ {}^w z_{l1} \end{bmatrix} = \begin{bmatrix} \frac{x_l}{z_l} \cos \theta_l - \sin \theta_l - \frac{d_i}{2} \\ \frac{x_l}{z_l} \sin \theta_l + \cos \theta_l \end{bmatrix} \quad (A7)$$

$${}^w P_{r1} = \begin{bmatrix} {}^w x_{r1} \\ {}^w z_{r1} \end{bmatrix} = \begin{bmatrix} \frac{x_r}{z_r} \cos \theta_r - \sin \theta_r + \frac{d_i}{2} \\ \frac{x_r}{z_r} \sin \theta_r + \cos \theta_r \end{bmatrix} \quad (A8)$$

According to ${}^w O_l = [-d_i/2, 0]^T$ and Equation (A7), the line ${}^l O_c P$ can be expressed as Equation (A9) in the world coordinate system $O_w X_w Z_w$.

$$z = \frac{{}^w z_{l1}}{\frac{d_i}{2} + {}^w x_{l1}} x + \frac{{}^w z_{r1} \times \frac{d_i}{2}}{\frac{d_i}{2} + {}^w x_{l1}} \quad (A9)$$

Based on ${}^w O_r = [d_i/2, 0]^T$ and Equation (A8), the line ${}^r O_c P$ in the world coordinate system $O_w X_w Z_w$ can be expressed as follows:

$$z = -\frac{{}^w z_{r1}}{\frac{d_i}{2} - {}^w x_{r1}} x + \frac{{}^w z_{r1} \times \frac{d_i}{2}}{\frac{d_i}{2} - {}^w x_{r1}} \quad (A10)$$

The intersection point of the lines ${}^r O_c P$ and ${}^l O_c P$ is the point P , so the depth z_w of point P can be obtained by Equations (A9) and (A10).

$$z_w = \frac{-d_i {}^w z_{r1} {}^w z_{r1}}{{}^w x_{r1} {}^w z_{l1} - {}^w z_{l1} \frac{d_i}{2} - {}^w z_{r1} \frac{d_i}{2} - {}^w x_{l1} {}^w z_{r1}} \quad (A11)$$

Appendix A.3. Measurement Error Analysis

In the real situation, there are the eyes' rotation angle errors (usually caused by time difference between image acquisition and motor angle acquisition when the two cameras move, the stepper motor's clearance error and the encoder's resolution) and image errors (usually caused by image distortion, image resolution and image feature extraction error). Let Δm_l and Δm_r be the image errors of the two cameras, respectively, then ${}^l P_{l1} = [x_l/z_l, 1]^T$ and ${}^r P_{r1} = [x_r/z_r, 1]^T$ can be revised as ${}^l P'_{l1} = [x_l/z_l + \Delta m_l, 1]^T$ and ${}^r P'_{r1} = [x_r/z_r + \Delta m_r, 1]^T$. Let $\Delta \theta_l$ and $\Delta \theta_r$ be the errors of two cameras' rotation angles, respectively. Therefore, we can rewrite Equations (A7) and (A8) as follows:

$${}^w P'_{l1} = \begin{bmatrix} {}^w x'_{l1} \\ {}^w z'_{l1} \end{bmatrix} = \begin{bmatrix} \left(\frac{x_l}{z_l} + \Delta m_l\right) \cos(\theta_l + \Delta \theta_l) - \sin(\theta_l + \Delta \theta_l) - \frac{d_i}{2} \\ \left(\frac{x_l}{z_l} + \Delta m_l\right) \sin(\theta_l + \Delta \theta_l) + \cos(\theta_l + \Delta \theta_l) \end{bmatrix} \quad (A12)$$

$${}^w P'_{r1} = \begin{bmatrix} {}^w x'_{r1} \\ {}^w z'_{r1} \end{bmatrix} = \begin{bmatrix} \left(\frac{x_r}{z_r} + \Delta m_r\right) \cos(\theta_r + \Delta \theta_r) - \sin(\theta_r + \Delta \theta_r) + \frac{d_i}{2} \\ \left(\frac{x_r}{z_r} + \Delta m_r\right) \sin(\theta_r + \Delta \theta_r) + \cos(\theta_r + \Delta \theta_r) \end{bmatrix} \quad (A13)$$

Based on the same principle, the revised depth z'_w of point P can be obtained as follows:

$$z'_w = \frac{-d_i {}^w z'_{l1} {}^w z'_{r1}}{{}^w x'_{r1} {}^w z'_{l1} - {}^w z'_{l1} \frac{d_i}{2} - {}^w z'_{r1} \frac{d_i}{2} - {}^w x'_{l1} {}^w z'_{r1}} \quad (\text{A14})$$

According to the cooperative movement pattern of human eyes, the absolute values of θ_l and θ_r are restricted to a limited range and assumed to be equal as follows:

$$-\theta_l = \theta_r = \theta \quad \text{s.t.} \quad 0 \leq \theta < \frac{\pi}{2} \quad (\text{A15})$$

Since Δm_r , Δm_l , $\Delta \theta_l$ and $\Delta \theta_r$ are usually close to 0, the simplified expression of the error Δz between the actual value z_w and measurement value z'_w can be derived from (A7), (A8) and (A11)–(A14).

$$\Delta z = z'_w - z_w \approx -z_w \frac{z_w (\Delta m_l \cos^2 \theta - \Delta m_r \cos^2 \theta - \Delta \theta_l + \Delta \theta_r)}{d_i + z_w (\Delta m_l \cos^2 \theta - \Delta m_r \cos^2 \theta - \Delta \theta_l + \Delta \theta_r)} \quad (\text{A16})$$

In addition, the relative error of z_w is

$$\Delta z_r = \frac{\Delta z}{z_w} \approx -\frac{z_w (\Delta m_l \cos^2 \theta - \Delta m_r \cos^2 \theta - \Delta \theta_l + \Delta \theta_r)}{d_i + z_w (\Delta m_l \cos^2 \theta - \Delta m_r \cos^2 \theta - \Delta \theta_l + \Delta \theta_r)} \quad (\text{A17})$$

Let

$$\varepsilon = \Delta m_l \cos^2 \theta - \Delta m_r \cos^2 \theta - \Delta \theta_l + \Delta \theta_r \quad (\text{A18})$$

In practice, ε has very small value and $z_w \varepsilon \ll d_i$, so Δz_r can be simplified as follows:

$$\Delta z_r \approx -\frac{z_w \varepsilon}{d_i} \quad (\text{A19})$$

From Equation (A19), it can be known that relative error of z_w is proportional to $z_w \varepsilon$ and inversely proportional to d_i . Thus, we can adopt the following strategies to reduce the error of depth:

- (1) Keep d_i long enough and constant when the bionic eyes move.
- (2) Observe the target as close as possible since the depth error is smaller when bionic eyes observe target in a close distance.
- (3) Control the two cameras of the bionic eyes with the same angular velocity during the process of the eyes' movement. In this way, $\Delta \theta^l$ and $\Delta \theta^r$ will be approximately equal to each other, and ε can be reduced.
- (4) Keep the target on the Z_w axis if possible, so that Δm_l and Δm_r are close to each other.

These strategies can be used to design effective motion control methods so that bionic eyes can perceive the target's 3D information accurately.

Appendix B. Proof of Equation (100)

Let

$$a = -d_i {}^w z'_{l1} {}^w z'_{r1} \quad (\text{A20})$$

$$b = {}^w x'_{r1} {}^w z'_{l1} - {}^w z'_{l1} \frac{d_i}{2} - {}^w z'_{r1} \frac{d_i}{2} - {}^w x'_{l1} {}^w z'_{r1} \quad (\text{A21})$$

Then, z'_w in Equation (A14) can be expressed as

$$z'_w = \frac{a}{b} \quad (\text{A22})$$

From Equations (A7) and (A8), we can obtain

$$w_{z_l1}' = \left(\frac{x_l}{z_l} + \Delta m_l\right) \sin(\theta_l + \Delta\theta_l) + \cos(\theta_l + \Delta\theta_l) \tag{A23}$$

$$w_{z_r1}' = \left(\frac{x_r}{z_r} + \Delta m_r\right) \sin(\theta_r + \Delta\theta_r) + \cos(\theta_r + \Delta\theta_r) \tag{A24}$$

From (A10), (A23) and (A24), we can obtain

$$\begin{aligned} a = & -d_i \frac{x_r}{z_r} \frac{x_l}{z_l} [\sin \theta_r \cos \Delta\theta_r + \cos \theta_r \sin \Delta\theta_r] [\sin \theta_l \cos \Delta\theta_l + \cos \theta_l \sin \Delta\theta_l] \\ & -d_i \frac{x_r}{z_r} [\sin \theta_r \cos \Delta\theta_r + \cos \theta_r \sin \Delta\theta_r] [\sin \theta_l \cos \Delta\theta_l + \cos \theta_l \sin \Delta\theta_l] \Delta m_l \\ & -d_i \frac{x_r}{z_r} [\sin \theta_r \cos \Delta\theta_r + \cos \theta_r \sin \Delta\theta_r] [\cos \theta_l \cos \Delta\theta_l - \sin \theta_l \sin \Delta\theta_l] \\ & -d_i \frac{x_l}{z_l} [\sin \theta_r \cos \Delta\theta_r + \cos \theta_r \sin \Delta\theta_r] [\sin \theta_l \cos \Delta\theta_l + \cos \theta_l \sin \Delta\theta_l] \Delta m_r \\ & -d_i [\sin \theta_r \cos \Delta\theta_r + \cos \theta_r \sin \Delta\theta_r] [\sin \theta_l \cos \Delta\theta_l + \cos \theta_l \sin \Delta\theta_l] \Delta m_l \Delta m_r \\ & -d_i [\sin \theta_r \cos \Delta\theta_r + \cos \theta_r \sin \Delta\theta_r] [\cos \theta_l \cos \Delta\theta_l - \sin \theta_l \sin \Delta\theta_l] \Delta m_r \\ & -d_i \frac{x_l}{z_l} [\sin \theta_l \cos \Delta\theta_l + \cos \theta_l \sin \Delta\theta_l] [\cos \theta_r \cos \Delta\theta_r - \sin \theta_r \sin \Delta\theta_r] \\ & -d_i [\sin \theta_l \cos \Delta\theta_l + \cos \theta_l \sin \Delta\theta_l] [\cos \theta_r \cos \Delta\theta_r - \sin \theta_r \sin \Delta\theta_r] \Delta m_l \\ & -d_i [\cos \theta_r \cos \Delta\theta_r - \sin \theta_r \sin \Delta\theta_r] [\cos \theta_l \cos \Delta\theta_l - \sin \theta_l \sin \Delta\theta_l] \end{aligned} \tag{A25}$$

$\Delta m_r, \Delta m_l, \Delta\theta_l$ and $\Delta\theta_r$ are usually close to 0, so

$$\begin{cases} \cos \Delta\theta_r \approx 1 \\ \cos \Delta\theta_l \approx 1 \\ \sin \Delta\theta_r \sin \Delta\theta_l \approx 0 \\ \Delta m_l \Delta m_r \approx 0 \\ \Delta m_l \sin \Delta\theta_l \approx 0 \\ \Delta m_r \sin \Delta\theta_l \approx 0 \\ \Delta m_l \sin \Delta\theta_r \approx 0 \\ \Delta m_r \sin \Delta\theta_r \approx 0 \end{cases} \tag{A26}$$

From (A25) and (A26), we can obtain

$$\begin{aligned} a \approx & d_i \frac{1}{z_r z_l} [-x_l x_r \sin \theta_l \sin \theta_r - x_r z_l \cos \theta_l \sin \theta_r - x_l z_r \sin \theta_l \cos \theta_r - z_r z_l \cos \theta_r \cos \theta_l \\ & + \sin \Delta\theta_l (-x_l x_r \sin \theta_r \cos \theta_l + x_r z_l \sin \theta_r \sin \theta_l - x_l z_r \cos \theta_r \cos \theta_l + z_r z_l \cos \theta_r \sin \theta_l) \\ & + \sin \Delta\theta_r (x_l z_r \sin \theta_l \sin \theta_r + z_r z_l \cos \theta_l \sin \theta_r - x_l x_r \sin \theta_l \cos \theta_r - x_r z_l \cos \theta_l \cos \theta_r) \\ & - \Delta m_l (x_r z_l \sin \theta_l \sin \theta_r + z_r z_l \sin \theta_l \cos \theta_r) - \Delta m_r (x_l z_r \sin \theta_r \sin \theta_l - z_r z_l \sin \theta_r \cos \theta_l)] \end{aligned} \tag{A27}$$

From (A7), (A8) and (A19), we can obtain

$$\begin{cases} x_l = (x_w + \frac{d_i}{2}) \cos \theta - z_w \sin \theta \\ z_l = -(x_w + \frac{d_i}{2}) \sin \theta + z_w \cos \theta \\ x_r = (x_w - \frac{d_i}{2}) \cos \theta + z_w \sin \theta \\ z_r = (x_w - \frac{d_i}{2}) \sin \theta + z_w \cos \theta \end{cases} \tag{A28}$$

Equation (A19) can be derived from (A19), (A27) and (A28):

$$\begin{aligned} a \approx & z_w^2 \frac{-d_i + [\frac{d_i}{2} \sin 2\theta + \frac{d_i(x_w + \frac{d_i}{2}) \sin^2 \theta}{z_w}]}{z_l z_r} \Delta m_l + z_w^2 \frac{[\frac{d_i(x_w - \frac{d_i}{2}) \sin^2 \theta}{z_w} - \frac{d_i}{2} \sin 2\theta] \Delta m_r}{z_l z_r} \\ & + z_w^2 \frac{-d_i(x_w + \frac{d_i}{2}) \sin \Delta\theta_l + \frac{d_i(x_w - \frac{d_i}{2}) \sin \Delta\theta_r}{z_w}}{z_l z_r} \end{aligned} \tag{A29}$$

$\Delta m_r, \Delta m_l, \Delta\theta_l$ and $\Delta\theta_r$ are usually close to 0, $x_w \ll z_w$, and $d_i \ll z_w$. Thus,

$$a \approx z_w^2 \frac{-d_i}{z_l z_r} \tag{A30}$$

From (A21), (A23) and (A24), we can obtain

$$\begin{aligned}
 b = & \frac{x_l}{z_l} \frac{x_r}{z_r} (\sin \theta_l \cos \Delta \theta_l + \cos \theta_l \sin \Delta \theta_l) (\cos \theta_r \cos \Delta \theta_r - \sin \theta_r \sin \Delta \theta_r) \\
 & + \frac{x_l}{z_l} (\sin \theta_l \cos \Delta \theta_l + \cos \theta_l \sin \Delta \theta_l) (\cos \theta_r \cos \Delta \theta_r - \sin \theta_r \sin \Delta \theta_r) \Delta m_r \\
 & - \frac{x_l}{z_l} (\sin \theta_l \cos \Delta \theta_l + \cos \theta_l \sin \Delta \theta_l) (\sin \theta_r \cos \Delta \theta_r + \cos \theta_r \sin \Delta \theta_r) \\
 & + \frac{x_r}{z_r} (\sin \theta_l \cos \Delta \theta_l + \cos \theta_l \sin \Delta \theta_l) (\cos \theta_r \cos \Delta \theta_r - \sin \theta_r \sin \Delta \theta_r) \Delta m_l \\
 & + (\sin \theta_l \cos \Delta \theta_l + \cos \theta_l \sin \Delta \theta_l) (\cos \theta_r \cos \Delta \theta_r - \sin \theta_r \sin \Delta \theta_r) \Delta m_l \Delta m_r \\
 & - (\sin \theta_l \cos \Delta \theta_l + \cos \theta_l \sin \Delta \theta_l) (\sin \theta_r \cos \Delta \theta_r + \cos \theta_r \sin \Delta \theta_r) \Delta m_l \\
 & + \frac{x_r}{z_r} (\cos \theta_l \cos \Delta \theta_l - \sin \theta_l \sin \Delta \theta_l) (\cos \theta_r \cos \Delta \theta_r - \sin \theta_r \sin \Delta \theta_r) \\
 & + (\cos \theta_l \cos \Delta \theta_l - \sin \theta_l \sin \Delta \theta_l) (\cos \theta_r \cos \Delta \theta_r - \sin \theta_r \sin \Delta \theta_r) \Delta m_r \\
 & - (\cos \theta_l \cos \Delta \theta_l - \sin \theta_l \sin \Delta \theta_l) (\sin \theta_r \cos \Delta \theta_r + \cos \theta_r \sin \Delta \theta_r) \\
 & - \frac{x_r}{z_r} \frac{x_l}{z_l} (\sin \theta_r \cos \Delta \theta_r + \cos \theta_r \sin \Delta \theta_r) (\cos \theta_l \cos \Delta \theta_l - \sin \theta_l \sin \Delta \theta_l) \\
 & - \frac{x_r}{z_r} (\sin \theta_r \cos \Delta \theta_r + \cos \theta_r \sin \Delta \theta_r) (\cos \theta_l \cos \Delta \theta_l - \sin \theta_l \sin \Delta \theta_l) \Delta m_l \\
 & + \frac{x_r}{z_r} (\sin \theta_r \cos \Delta \theta_r + \cos \theta_r \sin \Delta \theta_r) (\sin \theta_l \cos \Delta \theta_l + \cos \theta_l \sin \Delta \theta_l) \\
 & - \frac{x_l}{z_l} (\sin \theta_r \cos \Delta \theta_r + \cos \theta_r \sin \Delta \theta_r) (\cos \theta_l \cos \Delta \theta_l - \sin \theta_l \sin \Delta \theta_l) \Delta m_r \\
 & - (\sin \theta_r \cos \Delta \theta_r + \cos \theta_r \sin \Delta \theta_r) (\cos \theta_l \cos \Delta \theta_l - \sin \theta_l \sin \Delta \theta_l) \Delta m_l \Delta m_r \\
 & + (\sin \theta_r \cos \Delta \theta_r + \cos \theta_r \sin \Delta \theta_r) (\sin \theta_l \cos \Delta \theta_l + \cos \theta_l \sin \Delta \theta_l) \Delta m_r \\
 & - \frac{x_l}{z_l} (\cos \theta_r \cos \Delta \theta_r - \sin \theta_r \sin \Delta \theta_r) (\cos \theta_l \cos \Delta \theta_l - \sin \theta_l \sin \Delta \theta_l) \\
 & - (\cos \theta_r \cos \Delta \theta_r - \sin \theta_r \sin \Delta \theta_r) (\cos \theta_l \cos \Delta \theta_l - \sin \theta_l \sin \Delta \theta_l) \Delta m_l \\
 & + (\cos \theta_r \cos \Delta \theta_r - \sin \theta_r \sin \Delta \theta_r) (\sin \theta_l \cos \Delta \theta_l + \cos \theta_l \sin \Delta \theta_l)
 \end{aligned} \tag{A31}$$

From (A26) and (A31), we can obtain:

$$\begin{aligned}
 b \approx & \frac{1}{z_l z_r} [x_l x_r \sin \theta_l \cos \theta_r - x_l z_r \sin \theta_l \sin \theta_r + z_l x_r \cos \theta_l \cos \theta_r - z_l z_r \cos \theta_l \sin \theta_r \\
 & - x_l x_r \sin \theta_r \cos \theta_l + z_l z_r \sin \theta_l \cos \theta_r - x_l z_r \cos \theta_l \cos \theta_r + z_l x_r \sin \theta_l \sin \theta_r \\
 & + \sin \Delta \theta_l (x_l x_r \cos \theta_l \cos \theta_r - x_l z_r \cos \theta_l \sin \theta_r - z_l x_r \cos \theta_r \sin \theta_l + x_l z_r \sin \theta_l \cos \theta_r \\
 & + z_l z_r \cos \theta_r \cos \theta_l + z_l z_r \sin \theta_l \sin \theta_r + x_l x_r \sin \theta_l \sin \theta_r + z_l x_r \sin \theta_r \cos \theta_l) \\
 & + \sin \Delta \theta_r (-x_l x_r \sin \theta_l \sin \theta_r - x_l z_r \sin \theta_l \cos \theta_r - z_l x_r \cos \theta_l \sin \theta_r - z_l z_r \cos \theta_l \cos \theta_r \\
 & - x_l x_r \cos \theta_l \cos \theta_r + z_l x_r \sin \theta_l \cos \theta_r + x_l z_r \cos \theta_l \sin \theta_r - z_l z_r \sin \theta_l \sin \theta_r) \\
 & + \Delta m_l (z_l x_r \sin \theta_l \cos \theta_r - z_l z_r \sin \theta_l \sin \theta_r - z_l x_r \cos \theta_l \sin \theta_r - z_l z_r \cos \theta_l \cos \theta_r) \\
 & + \Delta m_r (x_l z_r \sin \theta_l \cos \theta_r + z_l z_r \cos \theta_l \cos \theta_r - x_l z_r \cos \theta_l \sin \theta_r + z_l z_r \sin \theta_l \sin \theta_r)]
 \end{aligned} \tag{A32}$$

Equation (A33) can be derived by (A19), (A28) and (A32):

$$\begin{aligned}
 b \approx & z_w \frac{-d_i - \Delta m_l [2x_w \sin \theta \cos \theta + \frac{(x_w^2 - d_l^2)}{z_w} \sin^2 \theta + z_w \cos^2 \theta]}{z_l z_r} \\
 & + z_w \frac{\Delta m_r [z_w \cos^2 \theta + \frac{(x_w^2 - d_l^2)}{z_w} \sin^2 \theta - 2x_w \sin \theta \cos \theta]}{z_l z_r} \\
 & + z_w \frac{\sin \Delta \theta_l [z_w + \frac{x_w^2 - d_l^2}{z_w}] - \sin \Delta \theta_r [z_w + \frac{x_w^2 - d_l^2}{z_w}]}{z_l z_r}
 \end{aligned} \tag{A33}$$

Δm_r , Δm_l , $\Delta \theta_l$ and $\Delta \theta_r$ are usually close to 0, $x_w \ll z_w$, and $d_i \ll z_w$. Thus,

$$b \approx z_w \frac{-d_i + z_w (\Delta m_r \cos^2 \theta - \Delta m_l \cos^2 \theta + \sin \Delta \theta_l - \sin \Delta \theta_r)}{z_l z_r} \tag{A34}$$

From (A22), (A29) and (A34), we can obtain

$$z'_w \approx \frac{z_w d_i}{d_i - z_w (\Delta m_r \cos^2 \theta - \Delta m_l \cos^2 \theta + \sin \Delta \theta_l - \sin \Delta \theta_r)} \tag{A35}$$

So,

$$\Delta z = z'_w - z_w \approx -z_w \frac{z_w (\Delta m_l \cos^2 \theta - \Delta m_r \cos^2 \theta - \sin \Delta \theta_l + \sin \Delta \theta_r)}{d_i + z_w (\Delta m_l \cos^2 \theta - \Delta m_r \cos^2 \theta - \sin \Delta \theta_l + \sin \Delta \theta_r)} \tag{A36}$$

$\Delta\theta_l$ and $\Delta\theta_r$ are usually close to 0, so

$$\Delta z \approx -z_w \frac{z_w(\Delta m_l \cos^2 \theta - \Delta m_r \cos^2 \theta - \Delta\theta_l + \Delta\theta_r)}{d_i + z_w(\Delta m_l \cos^2 \theta - \Delta m_r \cos^2 \theta - \Delta\theta_l + \Delta\theta_r)} \quad (\text{A37})$$

The proof is completed.

Appendix C. Two Equations Related to θ_t and θ_p

Substituting Equation (59) into Equation (60), we can obtain:

$$\begin{pmatrix} {}^1x_c \\ {}^1y_c \\ {}^1z_c \\ 1 \end{pmatrix} = \begin{pmatrix} {}^1n_x & {}^1o_x & {}^1a_x & {}^1p_x \\ {}^1n_y & {}^1o_y & {}^1a_y & {}^1p_y \\ {}^1n_z & {}^1o_z & {}^1a_z & {}^1p_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^1\theta_t & \sin^1\theta_t & 0 \\ 0 & -\sin^1\theta_t & \cos^1\theta_t & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos^1\theta_p & 0 & -\sin^1\theta_p & 0 \\ 0 & 1 & 0 & 0 \\ \sin^1\theta_p & 0 & \cos^1\theta_p & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix} \quad (\text{A38})$$

Equation (A38) can be factored out:

$$\begin{pmatrix} {}^1x_c \\ {}^1y_c \\ {}^1z_c \end{pmatrix} = \begin{pmatrix} {}^1n_x x_w \cos^1\theta_p - {}^1n_x z_w \sin^1\theta_p + {}^1o_x x_w \sin^1\theta_p \sin^1\theta_t \\ + {}^1o_x y_w \cos^1\theta_t + {}^1o_x z_w \cos^1\theta_p \sin^1\theta_t + {}^1a_x x_w \sin^1\theta_p \cos^1\theta_t \\ - {}^1a_x y_w \sin^1\theta_t + {}^1a_x z_w \cos^1\theta_p \cos^1\theta_t + {}^1p_x \\ {}^1n_y x_w \cos^1\theta_p - {}^1n_y z_w \sin^1\theta_p + {}^1o_y x_w \sin^1\theta_p \sin^1\theta_t \\ + {}^1o_y y_w \cos^1\theta_t + {}^1o_y z_w \cos^1\theta_p \sin^1\theta_t + {}^1a_y x_w \sin^1\theta_p \cos^1\theta_t \\ - {}^1a_y y_w \sin^1\theta_t + {}^1a_y z_w \cos^1\theta_p \cos^1\theta_t + {}^1p_y \\ {}^1n_z x_w \cos^1\theta_p - {}^1n_z z_w \sin^1\theta_p + {}^1o_z x_w \sin^1\theta_p \sin^1\theta_t \\ + {}^1o_z y_w \cos^1\theta_t + {}^1o_z z_w \cos^1\theta_p \sin^1\theta_t + {}^1a_z x_w \sin^1\theta_p \cos^1\theta_t \\ - {}^1a_z y_w \sin^1\theta_t + {}^1a_z z_w \cos^1\theta_p \cos^1\theta_t + {}^1p_z \end{pmatrix} \quad (\text{A39})$$

Substituting Equation (A39) into Equation (62), we can obtain:

$$\begin{pmatrix} \Delta u_l \\ \Delta v_l \end{pmatrix} = \begin{pmatrix} ({}^1k_x {}^1n_x x_w \cos^1\theta_p - {}^1k_x {}^1n_x z_w \sin^1\theta_p + {}^1k_x {}^1o_x x_w \sin^1\theta_p \sin^1\theta_t \\ + {}^1k_x {}^1o_x y_w \cos^1\theta_t + {}^1k_x {}^1o_x z_w \cos^1\theta_p \sin^1\theta_t + {}^1k_x {}^1a_x x_w \sin^1\theta_p \cos^1\theta_t \\ - {}^1k_x {}^1a_x y_w \sin^1\theta_t + {}^1k_x {}^1a_x z_w \cos^1\theta_p \cos^1\theta_t + {}^1k_x {}^1p_x) \\ \hline ({}^1n_z x_w \cos^1\theta_p - {}^1n_z z_w \sin^1\theta_p + {}^1o_z x_w \sin^1\theta_p \sin^1\theta_t \\ + {}^1o_z y_w \cos^1\theta_t + {}^1o_z z_w \cos^1\theta_p \sin^1\theta_t + {}^1a_z x_w \sin^1\theta_p \cos^1\theta_t \\ - {}^1a_z y_w \sin^1\theta_t + {}^1a_z z_w \cos^1\theta_p \cos^1\theta_t + {}^1p_z) \\ ({}^1k_y {}^1n_y x_w \cos^1\theta_p - {}^1k_y {}^1n_y z_w \sin^1\theta_p + {}^1k_y {}^1o_y x_w \sin^1\theta_p \sin^1\theta_t \\ + {}^1k_y {}^1o_y y_w \cos^1\theta_t + {}^1k_y {}^1o_y z_w \cos^1\theta_p \sin^1\theta_t + {}^1k_y {}^1a_y x_w \sin^1\theta_p \cos^1\theta_t \\ - {}^1k_y {}^1a_y y_w \sin^1\theta_t + {}^1k_y {}^1a_y z_w \cos^1\theta_p \cos^1\theta_t + {}^1k_y {}^1p_y) \\ \hline ({}^1n_z x_w \cos^1\theta_p - {}^1n_z z_w \sin^1\theta_p + {}^1o_z x_w \sin^1\theta_p \sin^1\theta_t \\ + {}^1o_z y_w \cos^1\theta_t + {}^1o_z z_w \cos^1\theta_p \sin^1\theta_t + {}^1a_z x_w \sin^1\theta_p \cos^1\theta_t \\ - {}^1a_z y_w \sin^1\theta_t + {}^1a_z z_w \cos^1\theta_p \cos^1\theta_t + {}^1p_z) \end{pmatrix} \quad (\text{A40})$$

Based on the same principle, substituting into each matrix and factoring the value of Δm_r , we can obtain

$$\begin{pmatrix} \Delta u_r \\ \Delta v_r \end{pmatrix} = \begin{pmatrix} \frac{({}^r k_x {}^r n_x {}^r x_w \cos {}^r \theta_p - {}^r k_x {}^r n_x {}^r z_w \sin {}^r \theta_p + {}^r k_x {}^r o_x {}^r x_w \sin {}^r \theta_p \sin {}^r \theta_t + {}^r k_x {}^r o_x {}^r y_w \cos {}^r \theta_t + {}^r k_x {}^r o_x {}^r z_w \cos {}^r \theta_p \sin {}^r \theta_t + {}^r k_x {}^r a_x {}^r x_w \sin {}^r \theta_p \cos {}^r \theta_t - {}^r k_x {}^r a_x {}^r y_w \sin {}^r \theta_t + {}^r k_x {}^r a_x {}^r z_w \cos {}^r \theta_p \cos {}^r \theta_t + {}^r k_x {}^r p_x)}{({}^r n_z {}^r x_w \cos {}^r \theta_p - {}^r n_z {}^r z_w \sin {}^r \theta_p + {}^r o_z {}^r x_w \sin {}^r \theta_p \sin {}^r \theta_t + {}^r o_z {}^r y_w \cos {}^r \theta_t + {}^r o_z {}^r z_w \cos {}^r \theta_p \sin {}^r \theta_t + {}^r a_z {}^r x_w \sin {}^r \theta_p \cos {}^r \theta_t - {}^r a_z {}^r y_w \sin {}^r \theta_t + {}^r a_z {}^r z_w \cos {}^r \theta_p \cos {}^r \theta_t + {}^r p_z)} \\ \frac{({}^r k_y {}^r n_y {}^r x_w \cos {}^r \theta_p - {}^r k_y {}^r n_y {}^r z_w \sin {}^r \theta_p + {}^r k_y {}^r o_y {}^r x_w \sin {}^r \theta_p \sin {}^r \theta_t + {}^r k_y {}^r o_y {}^r y_w \cos {}^r \theta_t + {}^r k_y {}^r o_y {}^r z_w \cos {}^r \theta_p \sin {}^r \theta_t + {}^r k_y {}^r a_y {}^r x_w \sin {}^r \theta_p \cos {}^r \theta_t - {}^r k_y {}^r a_y {}^r y_w \sin {}^r \theta_t + {}^r k_y {}^r a_y {}^r z_w \cos {}^r \theta_p \cos {}^r \theta_t + {}^r k_y {}^r p_y)}{({}^r n_z {}^r x_w \cos {}^r \theta_p - {}^r n_z {}^r z_w \sin {}^r \theta_p + {}^r o_z {}^r x_w \sin {}^r \theta_p \sin {}^r \theta_t + {}^r o_z {}^r y_w \cos {}^r \theta_t + {}^r o_z {}^r z_w \cos {}^r \theta_p \sin {}^r \theta_t + {}^r a_z {}^r x_w \sin {}^r \theta_p \cos {}^r \theta_t - {}^r a_z {}^r y_w \sin {}^r \theta_t + {}^r a_z {}^r z_w \cos {}^r \theta_p \cos {}^r \theta_t + {}^r p_z)} \end{pmatrix} \quad (A41)$$

By Equations (2), (A40) and (A41), Equation (A42) related to θ_t and θ_p can be obtained. It can be found from Equation (A32) that both θ_t and θ_p appear in the form of a trigonometric function, and it is difficult to obtain values of θ_t and θ_p directly from these two equations. In order to obtain the solution available in the project, we firstly obtain a sub-optimal observation pose and then use the sub-optimal observation pose as the initial value. We finally use the trial and error method to obtain the optimal observation pose.

$$\begin{pmatrix} ({}^1 k_x {}^1 n_x {}^1 x_w \cos \theta_p - {}^1 k_x {}^1 n_x {}^1 z_w \sin \theta_p + {}^1 k_x {}^1 o_x {}^1 x_w \sin \theta_p \sin \theta_t + {}^1 k_x {}^1 o_x {}^1 y_w \cos \theta_t + {}^1 k_x {}^1 o_x {}^1 z_w \cos \theta_p \sin \theta_t + {}^1 k_x {}^1 a_x {}^1 y_w \sin \theta_t + {}^1 k_x {}^1 a_x {}^1 x_w \sin \theta_p \cos \theta_t + {}^1 k_x {}^1 a_x {}^1 z_w \cos \theta_p \cos \theta_t + {}^1 k_x {}^1 p_x) \\ ({}^r n_z {}^r x_w \cos \theta_p - {}^r n_z {}^r z_w \sin \theta_p + {}^r o_z {}^r x_w \sin \theta_p \sin \theta_t + {}^r p_z + {}^r o_z {}^r y_w \cos \theta_t + {}^r o_z {}^r z_w \cos \theta_p \sin \theta_t + {}^r a_z {}^r x_w \sin \theta_p \cos \theta_t - {}^r a_z {}^r y_w \sin \theta_t + {}^r a_z {}^r z_w \cos \theta_p \cos \theta_t) = ({}^r k_x {}^r n_x {}^r x_w \cos \theta_p - {}^r k_x {}^r n_x {}^r z_w \sin \theta_p + {}^r k_x {}^r o_x {}^r x_w \sin \theta_p \sin \theta_t + {}^r k_x {}^r o_x {}^r y_w \cos \theta_t + {}^r k_x {}^r o_x {}^r z_w \cos \theta_p \sin \theta_t + {}^r k_x {}^r a_x {}^r x_w \sin \theta_p \cos \theta_t + {}^r k_x {}^r a_x {}^r y_w \sin \theta_t + {}^r k_x {}^r a_x {}^r z_w \cos \theta_p \cos \theta_t) ({}^1 n_z {}^1 x_w \cos \theta_p - {}^1 n_z {}^1 z_w \sin \theta_p + {}^1 o_z {}^1 x_w \sin \theta_p \sin \theta_t + {}^1 o_z {}^1 y_w \cos \theta_t + {}^1 o_z {}^1 z_w \cos \theta_p \sin \theta_t + {}^1 a_z {}^1 x_w \sin \theta_p \cos \theta_t - {}^1 a_z {}^1 y_w \sin \theta_t + {}^1 a_z {}^1 z_w \cos \theta_p \cos \theta_t + {}^1 p_z) \\ ({}^1 k_y {}^1 n_y {}^1 x_w \cos \theta_p - {}^1 k_y {}^1 n_y {}^1 z_w \sin \theta_p + {}^1 k_y {}^1 o_y {}^1 x_w \sin \theta_p \sin \theta_t + {}^1 k_y {}^1 o_y {}^1 y_w \cos \theta_t + {}^1 k_y {}^1 o_y {}^1 z_w \cos \theta_p \sin \theta_t - {}^1 k_y {}^1 a_y {}^1 y_w \sin \theta_t + {}^1 k_y {}^1 a_y {}^1 x_w \sin \theta_p \cos \theta_t + {}^1 k_y {}^1 a_y {}^1 z_w \cos \theta_p \cos \theta_t + {}^1 k_y {}^1 p_y) \\ ({}^r n_z {}^r x_w \cos \theta_p - {}^r n_z {}^r z_w \sin \theta_p + {}^r o_z {}^r x_w \sin \theta_p \sin \theta_t + {}^r p_z + {}^r o_z {}^r y_w \cos \theta_t + {}^r o_z {}^r z_w \cos \theta_p \sin \theta_t + {}^r a_z {}^r x_w \sin \theta_p \cos \theta_t - {}^r a_z {}^r y_w \sin \theta_t + {}^r a_z {}^r z_w \cos \theta_p \cos \theta_t) = ({}^r k_y {}^r n_y {}^r x_w \cos \theta_p - {}^r k_y {}^r n_y {}^r z_w \sin \theta_p + {}^r k_y {}^r o_y {}^r x_w \sin \theta_p \sin \theta_t + {}^r k_y {}^r o_y {}^r y_w \cos \theta_t + {}^r k_y {}^r o_y {}^r z_w \cos \theta_p \sin \theta_t + {}^r k_y {}^r a_y {}^r x_w \sin \theta_p \cos \theta_t + {}^r k_y {}^r a_y {}^r y_w \sin \theta_t + {}^r k_y {}^r a_y {}^r z_w \cos \theta_p \cos \theta_t) ({}^1 n_z {}^1 x_w \cos \theta_p - {}^1 n_z {}^1 z_w \sin \theta_p + {}^1 o_z {}^1 x_w \sin \theta_p \sin \theta_t + {}^1 o_z {}^1 y_w \cos \theta_t + {}^1 o_z {}^1 z_w \cos \theta_p \sin \theta_t + {}^1 a_z {}^1 x_w \sin \theta_p \cos \theta_t - {}^1 a_z {}^1 y_w \sin \theta_t + {}^1 a_z {}^1 z_w \cos \theta_p \cos \theta_t + {}^1 p_z) \end{pmatrix} \quad (A42)$$

References

1. Wang, Q.; Zou, W.; Xu, D.; Zhu, Z. Motion control in saccade and smooth pursuit for bionic eye based on three-dimensional coordinates. *J. Bionic Eng.* **2017**, *14*, 336–347. [\[CrossRef\]](#)
2. Kardamakis, A.A.; Moschovakis, A.K. Optimal control of gaze shifts. *J. Neurosci.* **2009**, *29*, 7723–7730. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Freedman, E.G.; Sparks, D.L. Coordination of the eyes and head: Movement kinematics. *Exp. Brain Res.* **2000**, *131*, 22–32. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Nakashima, R.; Fang, Y.; Hatori, Y.; Hiratani, A.; Matsumiya, K.; Kuriki, I.; Shioiri, S. Saliency-based gaze prediction based on head direction. *Vis. Res.* **2015**, *117*, 59–66. [\[CrossRef\]](#) [\[PubMed\]](#)
5. He, H.; Ge, S.S.; Zhang, Z. A saliency-driven robotic head with bio-inspired saccadic behaviors for social robotics. *Auton. Robot.* **2014**, *36*, 225–240. [\[CrossRef\]](#)
6. Law, J.; Shaw, P.; Lee, M. A biologically constrained architecture for developmental learning of eye-head gaze control on a humanoid robot. *Auton. Robot.* **2013**, *35*, 77–92. [\[CrossRef\]](#)

7. Wijayasinghe, I.B.; Aulisa, E.; Buttner, U.; Ghosh, B.K.; Glasauer, S.; Kremmyda, O. Potential and optimal target fixating control of the human head/eye complex. *IEEE Trans. Control Syst. Technol.* **2015**, *23*, 796–804. [[CrossRef](#)]
8. Ghosh, B.K.; Wijayasinghe, I.B.; Kahagalage, S.D. A geometric approach to head/eye control. *IEEE Access* **2014**, *2*, 316–332. [[CrossRef](#)]
9. Kuang, X.; Gibson, M.; Shi, B.E.; Rucci, M. Active vision during coordinated head/eye movements in a humanoid robot. *IEEE Trans. Robot.* **2012**, *28*, 1423–1430. [[CrossRef](#)]
10. Vannucci, L.; Cauli, N.; Falotico, E.; Bernardino, A.; Laschi, C. Adaptive visual pursuit involving eye-head coordination and prediction of the target motion. In Proceedings of the IEEE-RAS International Conference on Humanoid Robots, Madrid, Spain, 18–20 November 2014; pp. 541–546.
11. Hulse, M.; McBride, S.; Law, J.; Lee, M. Integration of active vision and reaching from a developmental robotics perspective. *IEEE Trans. Auton. Ment. Dev.* **2010**, *2*, 355–367. [[CrossRef](#)]
12. Anastasopoulos, D.; Naushahi, J.; Sklavos, S.; Bronstein, A.M. Fast gaze reorientations by combined movements of the eye, head, trunk and lower extremities. *Exp. Brain Res.* **2015**, *233*, 1639–1650. [[CrossRef](#)]
13. Daye, P.M.; Optican, L.M.; Blohm, G.; Lefèvre, P. Hierarchical control of two-dimensional gaze saccades. *J. Comput. Neurosci.* **2014**, *36*, 355–382. [[CrossRef](#)] [[PubMed](#)]
14. Rajruangrabin, J.; Popa, D.O. Robot head motion control with an emphasis on realism of neck–eye coordination during object tracking. *J. Intell. Robot. Syst.* **2011**, *63*, 163–190. [[CrossRef](#)]
15. Schulze, L.; Renneberg, B.; Lobmaier, J.S. Gaze perception in social anxiety and social anxiety disorder. *Front. Hum. Neurosci.* **2013**, *7*, 1–5. [[CrossRef](#)] [[PubMed](#)]
16. Liu, Y.; Zhu, D.; Peng, J.; Wang, X.; Wang, L.; Chen, L.; Li, J.; Zhang, X. Real-time robust stereo visual SLAM system based on bionic eyes. *IEEE Trans. Med. Robot. Bionics* **2020**, *2*, 391–398. [[CrossRef](#)]
17. Guitton, D. Control of eye–head coordination during orienting gaze shifts. *Trends Neurosci.* **1993**, *15*, 174–179. [[CrossRef](#)]
18. Matveev, A.S.; Hoy, M.C.; Savkin, A.V. 3D environmental extremum seeking navigation of a nonholonomic mobile robot. *Automatica* **2014**, *50*, 1802–1815. [[CrossRef](#)]
19. Nefti-Meziani, S.; Manzoor, U.; Davis, S.; Pupala, S.K. 3D perception from binocular vision for a low cost humanoid robot NAO. *Robot. Auton. Syst.* **2015**, *68*, 129–139. [[CrossRef](#)]
20. Surmann, H.; Nüchter, A.; Hertzberg, J. An autonomous mobile robot with a 3D laser range finder for 3D exploration and digitalization of indoor environments. *Robot. Auton. Syst.* **2003**, *45*, 181–198. [[CrossRef](#)]
21. Song, W.; Minami, M.; Shen, L.Y.; Zhang, Y.N. Bionic tracking method by hand & eye-vergence visual servoing. *Adv. Manuf.* **2016**, *4*, 157–166.
22. Li, H.Y.; Luo, J.; Huang, C.J.; Huang, Q.Z.; Xie, S.R. Design and control of 3-DoF spherical parallel mechanism robot eyes inspired by the binocular vestibule-ocular reflex. *J. Intell. Robot. Syst.* **2015**, *78*, 425–441. [[CrossRef](#)]
23. Masseck, O.A.; Hoffmann, K.P. Comparative neurobiology of the optokinetic reflex. *Ann. N. Y. Acad. Sci.* **2009**, *1164*, 430–439. [[CrossRef](#)] [[PubMed](#)]
24. Bruske, J.; Hansen, M.; Riehn, L.; Sommer, G. Biologically inspired calibration-free adaptive saccade control of a binocular camera-head. *Biol. Cybern.* **1997**, *77*, 433–446. [[CrossRef](#)]
25. Wang, X.; Van De Weem, J.; Jonker, P. An advanced active vision system imitating human eye movements. In Proceedings of the 2013 16th International Conference on Advanced Robotics, Montevideo, Uruguay, 25–29 November 2013; pp. 5–10.
26. Antonelli, M.; Duran, A.J.; Chinellato, E.; Pobil, A.P. Adaptive saccade controller inspired by the primates’ cerebellum. In Proceedings of the IEEE International Conference on Robotics and Automation, Seattle, WA, USA, 26–30 May 2015; pp. 5048–5053.
27. Robinson, D.A.; Gordon, J.L.; Gordon, S.E. A model of the smooth pursuit eye movement system. *Biol. Cybern.* **1986**, *55*, 43–57. [[CrossRef](#)] [[PubMed](#)]
28. Brown, C. Gaze controls with interactions and delays. *IEEE Trans. Syst. Man Cybern.* **1990**, *20*, 518–527. [[CrossRef](#)]
29. Deno, D.C.; Keller, E.L.; Crandall, W.F. Dynamical neural network organization of the visual pursuit system. *IEEE Trans. Biomed. Eng.* **1989**, *36*, 85–92. [[CrossRef](#)] [[PubMed](#)]
30. Lunghi, F.; Lazzari, S.; Magenes, G. Neural adaptive predictor for visual tracking system. In Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Hong Kong, China, 1 November 1998; Volume 20, pp. 1389–1392.
31. Lee, W.J.; Galiana, H.L. An internally switched model of ocular tracking with prediction. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2005**, *13*, 186–193. [[CrossRef](#)]
32. Avni, O.; Borrelli, F.; Katzir, G.; Rivlin, E.; Rotstein, H. Scanning and tracking with independent cameras—a biologically motivated approach based on model predictive control. *Auton. Robot.* **2008**, *24*, 285–302. [[CrossRef](#)]
33. Zhang, M.; Ma, X.; Qin, B.; Wang, G.; Guo, Y.; Xu, Z.; Wang, Y.; Li, Y. Information fusion control with time delay for smooth pursuit eye movement. *Physiol. Rep.* **2016**, *4*, e12775. [[CrossRef](#)]
34. Santini, F.; Rucci, M. Active estimation of distance in a robotic system that replicates human eye movement. *Robot. Auton. Syst.* **2007**, *55*, 107–121. [[CrossRef](#)]
35. Chinellato, E.; Antonelli, M.; Grzyb, B.J.; Del Pobil, A.P. Implicit sensorimotor mapping of the peripersonal space by gazing and reaching. *IEEE Trans. Auton. Ment. Dev.* **2011**, *3*, 43–53. [[CrossRef](#)]

36. Song, Y.; Zhang, X. An active binocular integrated system for intelligent robot vision. In Proceedings of the IEEE International Conference on Intelligence and Security Informatics, Washington, DC, USA, 11–14 June 2012; pp. 48–53.
37. Wang, Y.; Zhang, G.; Lang, H.; Zuo, B.; De Silva, C.W. A modified image-based visual servo controller with hybrid camera configuration for robust robotic grasping. *Robot. Auton. Syst.* **2014**, *62*, 1398–1407. [[CrossRef](#)]
38. Lee, Y.C.; Lan, C.C.; Chu, C.Y.; Lai, C.M.; Chen, Y.J. A pan-tilt orienting mechanism with parallel axes of flexural actuation. *IEEE-ASME Trans. Mechatron.* **2013**, *18*, 1100–1112. [[CrossRef](#)]
39. Wang, Q.; Zou, W.; Zhang, F.; Xu, D. Binocular initial location and extrinsic parameters real-time calculation for bionic eye system. In Proceedings of the 11th World Congress on Intelligent Control and Automation, Shenyang, China, 29 June–4 July 2014; pp. 74–80.
40. Fan, D.; Liu, Y.Y.; Chen, X.P.; Meng, F.; Liu, X.L.; Ullah, Z.; Cheng, W.; Liu, Y.H.; Huang, Q. Eye gaze based 3D triangulation for robotic bionic eyes. *Sensors* **2020**, *20*, 5271. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.