Article

The Design and Control of a Biomimetic Binocular Cooperative Perception System Inspired by the Eye Gaze Mechanism

Xufang Qin [1], Xiaohua Xia [1],* , Zhaokai Ge [1], Yanhao Liu [2] and Pengju Yue [1]

[1] Key Laboratory of Road Construction Technology and Equipment of MOE, Chang'an University, Xi'an 710064, China
[2] TianQin Research Center for Gravitational Physics and School of Physics and Astronomy, Sun Yat-sen University (Zhuhai Campus), Zhuhai 519082, China
* Correspondence: xhxia@chd.edu.cn

Abstract: Research on systems that imitate the gaze function of human eyes is valuable for the development of humanoid eye intelligent perception. However, the existing systems have some limitations, including the redundancy of servo motors, a lack of camera position adjustment components, and the absence of interest-point-driven binocular cooperative motion-control strategies. In response to these challenges, a novel biomimetic binocular cooperative perception system (BBCPS) was designed and its control was realized. Inspired by the gaze mechanism of human eyes, we designed a simple and flexible biomimetic binocular cooperative perception device (BBCPD). Based on a dynamic analysis, the BBCPD was assembled according to the principle of symmetrical distribution around the center. This enhances braking performance and reduces operating energy consumption, as evidenced by the simulation results. Moreover, we crafted an initial position calibration technique that allows for the calibration and adjustment of the camera pose and servo motor zero-position, to ensure that the state of the BBCPD matches the subsequent control method. Following this, a control method for the BBCPS was developed, combining interest point detection with a motion-control strategy. Specifically, we propose a binocular interest-point extraction method based on frequency-tuned and template-matching algorithms for perceiving interest points. To move an interest point to a principal point, we present a binocular cooperative motion-control strategy. The rotation angles of servo motors were calculated based on the pixel difference between the principal point and the interest point, and PID-controlled servo motors were driven in parallel. Finally, real experiments validated the control performance of the BBCPS, demonstrating that the gaze error was less than three pixels.

Keywords: gaze mechanism; bionic binocular visual system; binocular cooperative motion

1. Introduction

The gaze function is a crucial biological feature of the human visual system. It enables human eyes to identify interest targets in the environment and swiftly shift the gaze to these targets, placing the targets in the foveal region. By doing this, humans can obtain more details and information about the interest targets in the environment and less information about uninteresting regions [1–3]. Imitating the gaze function holds great significance for the advancement of fields such as human–robot interaction [4,5], autonomous driving [6], virtual reality [7], etc. Moreover, imitating the gaze function for the humanoid eye perception system has the potential to filter redundant information from huge data, optimize the use of computing and storage resources, enhance scene comprehension, and improve perception accuracy. This imitation stands as an important step in advancing humanoid eye intelligent perception.

The primary work in imitating the eye gaze function is to research the eye gaze mechanism. The mechanism of eye movement, which plays a key role in the gaze function, has been widely studied. Marg introduced electro-oculography (EOG) as a method

for measuring eye movement by obtaining eye potentials through electrodes around the eyes [8]. However, this contact measurement method suffers from lower precision and poor portability. Subsequently, the presentation of video-oculography (VOG) offered a more accurate and portable non-contact eye movement measurement method [9]. VOG used cameras mounted on wearable devices to capture the position of the pupil. Nevertheless, this method lacked a stimulus-presenting monitor and recording device, preventing independent measurements. The design of the all-in-one eye-movement-measuring device overcame this limitation, enabling independent and accurate eye movement measurement [10]. Through the gradual improvement of eye-movement-measurement devices, the factors affecting the eye movement mechanism have been studied [11–14]. Those studies have found that factors such as gender differences [12], cross-cultural differences in facial features [13], and stimulus contrast and spatial position differences [14] contribute to saccadic pattern differences. Now, many researchers are interested applying our increasingly robust understanding of the eye movement mechanism to the structural design and motion control of humanoid eye perception systems, which can imitate the gaze function.

In this structural design field, the methods can be divided into two categories. The first method refers to imitating the physiological structure of the extraocular muscles. This usually involves using a spherical link structure [15], spherical parallel mechanism [16], or multiple flexible ropes parallel mechanism [17] to design a device that achieves a highly realistic imitation of the physiological structure of the human eye. However, researchers face difficulties in reducing the sizes of these devices. With that goal, some studies have proposed that super-coiled polymers (SCPs) [18] or pneumatic artificial muscles (PAMs) [19] can be used to replace rigid materials or ropes in the design of a device. However, achieving precise control of these devices has remained challenging. The second method is to imitate the effect of the actual motion of the eye. This method usually uses the servo motor as the power source, which can reduce difficulty of controlling the device. Fan et al. [20] designed a bionic eye that could tilt, pan, and roll with six servo motors. However, a rolling motion is generally not required for a device that imitates the gaze function [21]. Some studies have focused on the design of devices that can tilt and pan with four servo motors [22,23]. However, those four servo motors are not synced for the cooperative motion of human eyes. Thus, a device was designed that could tilt and pan with three servo motors [24]. However, this device lacked a camera position adjustment component. It would have been difficult to ensure that the vertical visual field of the two cameras was consistent due to the potential for assembly errors, which affect gaze accuracy. Moreover, in a dynamic analysis [25,26] of this device, it was found that the torque of the servo motor responsible for tilting exceeded that of the other two servo motors, impacting the overall performance and efficiency. Therefore, further optimization of the structure is needed.

In regard to motion control, researchers have explored two distinct approaches. One approach is a motion-control strategy driven by a set target [22,27–31]. Mao et al. [28] proposed a control method that could be described as a two-level hierarchical system. This method could imitate the horizontal saccade of human eyes. Subsequently, a control method [30] was designed that employed a hierarchical neural network model based on predictive coding/biased competition with divisive input modulation (PC/BC-DIM) to achieve gaze shifts in 3D space. Despite the effect of the control method being consistent with that of human eye movement, the neural network requires plenty of data support compared to traditional algorithms. As an alternative, a traditional control algorithm based on 3D or 2D information of a target has raised attention [22,31,32]. For example, a vision servo based on 2D images was proposed to control the pose of each camera and achieve the fixation of a target [22]. Rubies et al. [32] calibrated the relationship between the eye positions on a robot's iconic face displayed on a flat screen and the 3D coordinates of a target point, thereby controlling the gaze toward the target. The motion-control strategy driven by the set target has clear objectives and facilitates precise motion control. However, it falls short in imitating the spontaneous gaze behavior of humans and may ignore other key information, which leaves it impossible to fully understand the scene. Another approach is

a motion-control strategy driven by a salient point, which can make up for these shortages. Researchers have made significant progress in saliency-detection algorithms, including classical algorithms [33–35] and deep neural networks [36–38]. The results of saliency detection are increasingly aligning with human eye selective attention. Building on this foundation, Zhu et al. introduced a saccade control strategy driven by binocular attention based on information maximization (BAIM) and a convergence control strategy based on a two-layer neural network [39]. However, this method cannot simultaneously execute saccade control and convergence control to imitate the cooperative motion of human eyes.

Recognizing the limitations of the above work, we have proposed a design and control method of the BBCPS inspired by the gaze mechanism of human eyes. To address the issues of servo motor redundancy and the lack of camera position adjustment components in the existing systems, a simple and flexible BBCPD was designed. The BBCPD consisted of RGB cameras, servo motors, pose adjustment modules, braced frames, calibration objects, transmission frames, and bases. It was assembled according to the innovative principle of symmetrical distribution around the center. A simulation demonstrated that the BBCPD achieved a great reduction in energy consumption and an enhancement in braking response performance. Furthermore, we developed an initial position calibration technique to ensure that the state of the BBCPD could meet the requirement of the subsequent control method. On this basis, we propose a control method of the BBCPS, aiming to fill the gap in binocular cooperative motion-control strategies, driven by interest points in the existing systems. In the proposed control method, a PID controller is introduced to realize precise control of a single servo motor. A binocular interest-point extraction method based on frequency-tuned and template-matching algorithms is presented to identify interest points. A binocular cooperative motion-control strategy is then outlined to coordinate the motion of servo motors in order to move the interest point to the principal point. Finally, we summarize the results of real experiments, which proved that the control method of the BBCPS could control the gaze error within three pixels.

The main contributions of our work are as follows. (a) We designed and controlled the BBCPS to simulate the human eye gaze function. This contributes to deepening our understanding of human eye gaze mechanisms and advancing the field of humanoid eye intelligent perception. (b) Our designed BBCPD features a simple structure, flexibility and adjustability, low energy consumption, and excellent braking performance. (c) We developed an interest-point-driven binocular cooperative motion-control method, perfecting the research on the control strategy for imitating human eye gaze. Additionally, we calibrated the initial position of the BBCPS via our self-developed calibration technique. This eliminates the need for repeated calibration in subsequent applications, improving the operational convenience of the BBCPS. What's more, our proposed binocular interest-point extraction method based on frequency-tuned and template-matching algorithms enriches the current research in the field of salient point detection.
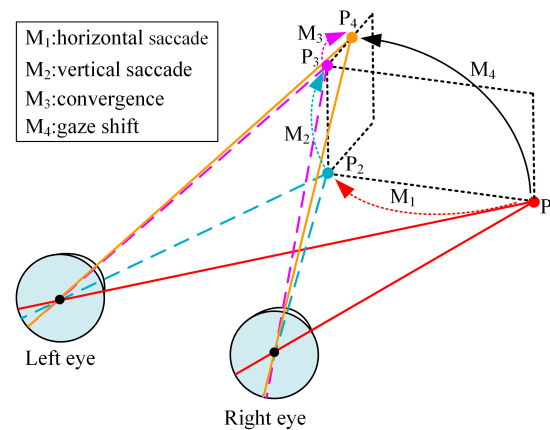
2. Gaze Mechanisms of Human Eyes

The movement of the eyeball is crucial to the gaze function. As shown in Figure 1, the eyeball is usually regarded as a perfect sphere, and its movement is controlled by the medial rectus, lateral rectus, superior rectus, inferior rectus, superior oblique, and inferior oblique [40–42]. These muscles contract and relax to perform different eye movements. The superior and inferior oblique muscles assist in the torsional movement of the eyeball. Torsional eye movements, characterized by minimal overall variability (approximately 0.10°), are an unconscious reflex and strictly physiologically controlled [21]. The superior and inferior rectus muscles rotate the eyeball around the horizontal axis, and the lateral and medial rectus muscles allow the eyeball to rotate around the vertical axis.
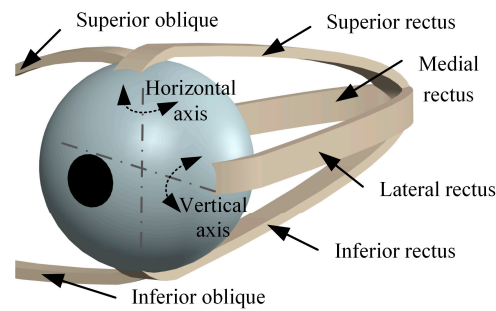
Figure 1. Eyeball movement.

Vertical and horizontal movements of the eyeball are important for the line-of-sight shift in the gaze function [43], which refers to the process of shifting the current line of sight to the interest point through eyeball movements during visual observation. This process involves saccade and convergence. Saccade is a conjugate movement that can achieve the line-of-sight shift of human eyes in both horizontal and vertical directions. Convergence describes a non-conjugate movement of human eyes in the horizontal direction, where the two eyes move in opposite directions to help humans observe points at different depths. By coordinating saccade and convergence, the two eyes can shift their line of sight to any point of interest in three-dimensional space.

To better understand the movement mechanism of the gaze function, we have created a schematic diagram of the human eye cooperative movement, shown in Figure 2. From a physiological point of view, the human eye changes from the gaze point P_1 to the gaze point P_4 through the coordination of saccade and convergence. We assume that the eye movement is sequential, and the process M_4 is decomposable into M_1, M_2, and M_3. Specifically, the shift from the gaze point P_1 to the gaze point P_2 is first achieved through the horizontal saccadic movement M_1. The shift from the gaze point P_2 to the gaze point P_3 is then accomplished through the vertical saccadic movement M_2. In the end, the convergent movement M_3 is employed to shift the gaze point P_3 to the gaze point P_4.

**Figure 2.** Human eye cooperative movement.

3. Structural Design

Inspired by the gaze mechanism of human eyes, the mechanical structure of the BBCPD was designed. Its 3D model is shown in Figure 3. The device is composed of two RGB cameras, three servo motors, two pose adjustment modules, two braced frames, two calibration objects, a transmission frame, and a base. RGB cameras capture images, and servo motors act as the power source. Pose adjustment modules are used to accurately adjust the pose of cameras toward different desired locations. This indicates an increased flexibility of the BBCPD. The transmission frame is designed to transmit motion. Calibration objects are used to calibrate the initial position of the BBCPD. Braced frames serve to ensure

the suspension of the transmission frame, guaranteeing the normal operation of the upper servo motor. The role of the base is to ensure the stable operation of the BBCPD.
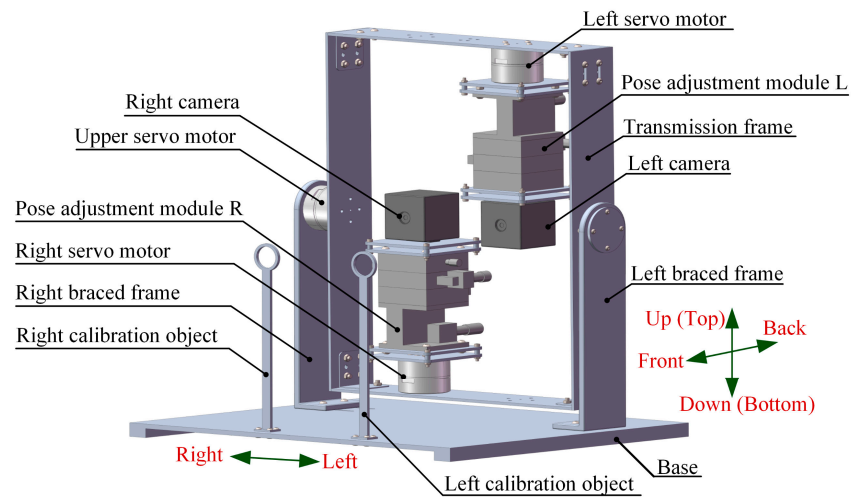

Figure 3. Three-dimensional model of the BBCPD.

In the component design, the transmission frame and the base are designed as parallel symmetrical structures motivated by the stability of the symmetrical structure. The braced frame adopts an L-shaped structure because this is highly stable and its different ends can be used to connect various other components. The design inspiration for the pose adjustment module is derived from the screw motion mechanism and the turbine worm drive. Based on the former, the camera can be adjusted in three directions: front–back, left–right, and up–down. Simultaneously, inspired by the turbine worm drive to change the direction of rotation, the roll, pan, and tilt adjustment of the camera are skillfully realized. The top of the calibration object is designed as a thin-walled ring, as the circle center is easy to detect, facilitating the subsequent zero-position adjustment of the servo motor. Furthermore, lightweight and high-strength aluminum alloy is selected as the component material.

During the assembly of the components, the principle of symmetrical distribution about the center of the transmission frame is followed, though it loses some of its biomimetic morphology compared to the classic principle of a symmetrical low center of gravity [2]. We recognize that the torque required of the upper servo motor in the BBCPD is markedly smaller than that in a system assembled according to the symmetrical low center of gravity principle. This means that the power consumption is lower in the BBCPD. In addition, the rotational inertia of the upper servo motor load around the rotation axis in the BBCPD is smaller than that in the system assembled according to the symmetrical low center of gravity principle. The BBCPD also has better braking performance. The detailed reason will be explained when we present our subsequent dynamic analysis of the upper servo motor.

In Figure 4, c_t is the center of mass of the whole load, and D represents the vertical distance from c_t to the rotation axis. The rotation axis serves as the boundary, and the load is divided into upper and lower parts. The center of mass of the upper part is denoted as c_u, and the vertical distance from c_u to the rotation axis is defined as r_u. The symbol c_d is the center of mass of the lower part, and the vertical distance from c_d to the rotation axis is expressed as r_d. L_i ($i = 0, 1, \dots$) represents the location of the motor, and θ_i denotes the rotation angle of the motor at L_i. The mass of the upper part is represented as m_u, m_d represents the mass of the lower part, and g means the acceleration of gravity.
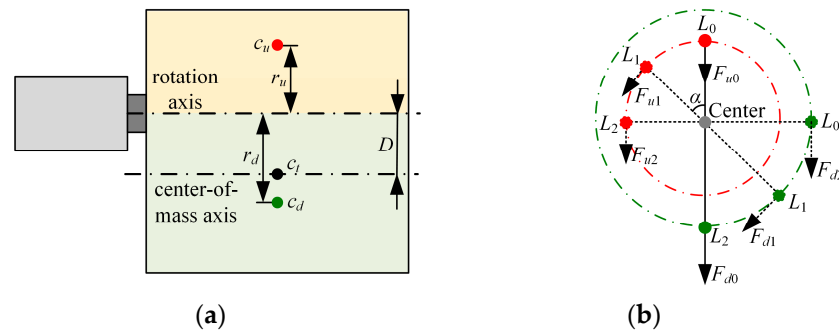
Figure 4. Dynamic analysis. (**a**) Structural sketch of the upper servo motor and its load, (**b**) force analysis of the upper servo motor.

According to the parallel-axis theorem, the rotational inertia of the whole load around the rotation axis J is given by Equation (1).

$$J = J_c + mD^2, \tag{1}$$

where m represents the mass of the load of the upper servo motor, and J_c is the rotational inertia of the whole load around the center-of-mass axis.

Next, we conducted the force analysis on the motor and derived the torque of the load on the motor at L_i.

$$M_i = F_{ui} r_u - F_{di} r_d = m_u g \sin \theta_i r_u - m_d g \sin \theta_i r_d. \tag{2}$$

Thus, the torque of the motor at L_i could be obtained by combining Equations (1) and (2).

$$T_i = J\omega - M_i = (J_c + mD^2)\omega - (m_u g \sin \theta_i r_u - m_d g \sin \theta_i r_d), \tag{3}$$

where ω is the angular acceleration of the motor. According to Equation (3), it can be observed that as D and the difference between m_u and m_d decrease, T_i and J become smaller. In the BBCPD, D and the difference between m_u and m_d are close to 0. However, in the system assembled according to the symmetrical low center of gravity principle, both D and the difference between m_u and m_d are greater than 0.

Afterward, we performed stress–strain analyses on the BBCPD using software and refined the dimensions of the components. The initial position of the BBCPD was defined as the state that imitates the approximately symmetrical distribution of eyeballs about the midline of the face when humans gaze at infinity. In other words, the camera optical center coincides with the rotation center, and the cameras are parallel to each other. Ultimately, the layout of each component in the space is introduced in the order from left to right.

The bottom of the left servo motor is installed at the bottom of the transmission frame, 64 mm to the left side of the transmission frame. Its shaft end is connected to the bottom of the pose adjustment module L, and the top of the pose adjustment module L is linked to the bottom of the left camera. The optical center of the left camera passes through the shaft of the left servo motor. The bottom of the right servo motor is installed at the top of the transmission frame, 64 mm to the right side of the transmission frame. Its shaft end is connected to the bottom of the pose adjustment module R, and the top of the pose adjustment module R is linked to the base of the right camera. The optical center of the right camera passes the shaft of the right servo motor. The shaft end of the upper servo motor is connected to the right side of the transmission frame, 180 mm above the bottom of the transmission frame, and the shaft passes through the optical centers of two cameras. The bottom of the upper servo motor is connected to the top of the braced frame. Braced frames are installed on the left and right sides of the base in opposite poses to make full use of the space. The left and right calibration objects are vertically fixed to the front of the

base, located at 125 mm and 245 mm on the left side of the base, respectively. The plane they lie on is parallel to the planes containing the shafts of the three servo motors.

The BBCPD with three degrees of freedom can effectively imitate the cooperative motion of human eyes. Specifically, the left servo motor and the right servo motor drive the left camera and the right camera to pan, respectively, thereby realizing the imitation of horizontal saccade and convergence. The upper servo motor drives the left camera and the right camera to tilt simultaneously through the transmission frame to imitate the vertical saccade. The design fully considers the human eye gaze mechanism and provides hardware support for imitating the gaze function of human eyes.

4. Initial Position Calibration

The initial position calibration of the BBCPD is a crucial step for achieving the control of the BBCPS. This is designed so that the initial position of the BBCPD meets the requirements of the subsequent control method that the camera optical center coincides with the rotation center and the cameras are parallel to each other. In addition, once the initial position is determined, the BBCPD does not need to be recalibrated during subsequent applications, saving time and resources. In the design of the BBCPD, we default that the initial position of the BBCPD is in line with the requirement of the subsequent control method. However, due to inevitable errors during the manufacturing and assembly processes, the initial position of the real BBCPD makes it challenging to guarantee this requirement is met. Additionally, the zero-positions of servo motors may not be set at the ideal initial position for the real BBCPD. Therefore, we provided the initial position calibration technology of the BBCPD to determine the initial position by calibrating and adjusting the camera poses and the zero-positions of servo motors.

4.1. Camera Pose Calibration and Adjustment

Camera pose calibration and adjustment refers to calibrating the rotation and translation parameters from the base coordinate frame to the camera coordinate frame, and then changing the camera pose using the pose adjustment module. The camera coordinate frame is a right-handed coordinate frame with the optical center as the origin and straight lines parallel to the length and width of the photosensitive plane as the horizontal axis and the vertical axis. The base coordinate frame is defined as a right-handed coordinate frame with the rotation center as the origin point, the horizontal rotation axis as the horizontal axis, and the vertical rotation axis as the vertical axis.

First, we devoted our time to calibrating the rotation parameters, in order to adjust the camera poses to achieve mutual parallelism of the photosensitive planes of the two cameras. Taking the left camera as an example, under corrected lens distortion, the calibration principle of the rotation parameters is shown in Figure 5. The camera coordinate frame is denoted as $O_c X_c Y_c Z_c$. The base coordinate frame is represented by $O_b X_b Y_b Z_b$, and the calibration object coordinate frame is described by $O_o X_o Y_o Z_o$. The original point O_o is the ring center of the left calibration object, and the plane $O_o X_o Y_o$ is the vertical center plane of the ring. $O_w X_w Y_w Z_w$ is the world coordinate frame. The original point O_w is set at the upper-right corner of the checkerboard. The horizontal and vertical directions of the checkerboard are the directions of the axis X_w and the axis Y_w, respectively.

The checkerboard is placed against the calibration objects so that the plane $O_w X_w Y_w$ is parallel to the plane $O_o X_o Y_o$. In the real BBCPD, the plane $O_o X_o Y_o$ is parallel to the plane $O_b X_b Y_b$. Thus, by deriving the rotation relationship between the coordinate frame $O_w X_w Y_w Z_w$ and the coordinate frame $O_c X_c Y_c Z_c$, the rotation parameters can be calibrated. According to the linear projection principle [44], the relationship between the 3D corner point P_i and the 2D corner point p_i is expressed via Equation (4).

$$p_i = sK\begin{bmatrix} R & t \end{bmatrix} P_i, \tag{4}$$

where K is the intrinsic matrix, s is the scale factor, R is the rotation matrix, and t is the translation matrix. We use all corners of the checkerboard to solve R with known P_i, p_i,

and K. Based on Equation (4), Equation (5) is obtained, and then R is calculated using the least-squares method.

$$K^{-1}[p_0, \, p_1, \, \cdots, \, p_n] = s \begin{bmatrix} R & t \end{bmatrix} [P_0, \, P_1, \, \cdots, \, P_n], \tag{5}$$

where n ($n > 4$) is the number of corners. Due to the small errors caused by the manufacturing and assembly processes in normal cases, the absolute values of the parameters θ_x, θ_y, and θ_z are generally no more than $90°$. The rotation parameters θ_x, θ_y, and θ_z can be uniquely calculated by using Equation (6) based on the rotation order of $Z_w Y_w X_w$.

$$\begin{cases} \theta_x = \arctan \frac{-R_{23}}{R_{33}} \\ \theta_y = \arctan \frac{R_{13}}{-\sin\theta_x R_{23} + \cos\theta_x R_{33}} \\ \theta_z = \arctan \frac{\cos\theta_x R_{21} + \sin\theta_x R_{31}}{\cos\theta_x R_{22} + \sin\theta_x R_{32}} \end{cases}, \text{ with } R^{-1} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}. \tag{6}$$
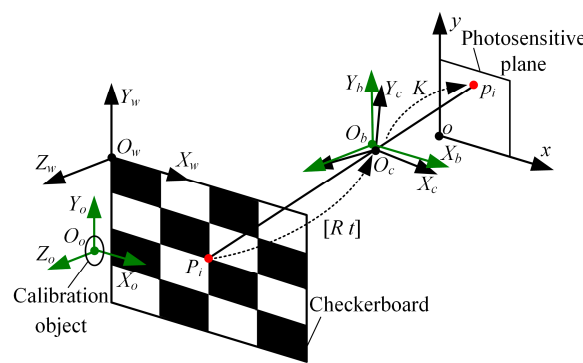

Figure 5. Calibration principle of the rotation parameters.

According to the above principle, the rotation parameters of the right camera can be solved. The poses of the left and right cameras are changed by sequentially adjusting the tilting, panning, and rolling angles of the pose adjustment module L and the pose adjustment module R. The tilting, panning, and rolling angles are the corresponding negative rotation parameters. At this time, the camera coordinate frame is parallel to the base coordinate frame, and the photosensitive planes are parallel to each other.

Next, we calibrated the translation parameters to adjust the camera poses so that the camera optical center coincided with the rotation center [45]. Based on the above steps, the calibration principle of the translation parameters, taking the left camera as an example, is depicted in Figure 6. The camera coordinate frame at the initial position $O_{c0}X_{c0}Y_{c0}Z_{c0}$ is parallel to the base coordinate frame $O_b X_b Y_b Z_b$. $O_{c1}X_{c1}Y_{c1}Z_{c1}$ represents the camera coordinate frame after motion. The end-of-motion coordinate frame at the initial position $O_{e0}X_{e0}Y_{e0}Z_{e0}$ is established by taking the line connecting the point O_{c0} and the point O_b as the X_{e0}-axis, the Y_b-axis as the Y_{e0}-axis, and the point O_b as the origin. The end-of-motion coordinate frame after motion is denoted as $O_{e1}X_{e1}Y_{e1}Z_{e1}$. The rotation angles of the camera around the Y_b-axis and X_b-axis are θ and β. The translation parameters in X_b, Y_b, and Z_b are denoted as l_x, l_y, and l_z.

Let T_{ij} represent the transformation from the coordinate frame $O_i X_i Y_i Z_i$ to the coordinate frame $O_j X_j Y_j Z_j$. Based on the transformation relationship between the coordinate frame $O_{c0}X_{c0}Y_{c0}Z_{c0}$ and the coordinate frame $O_{c1}X_{c1}Y_{c1}Z_{c1}$ in Figure 6a, Equation (7) can be obtained.
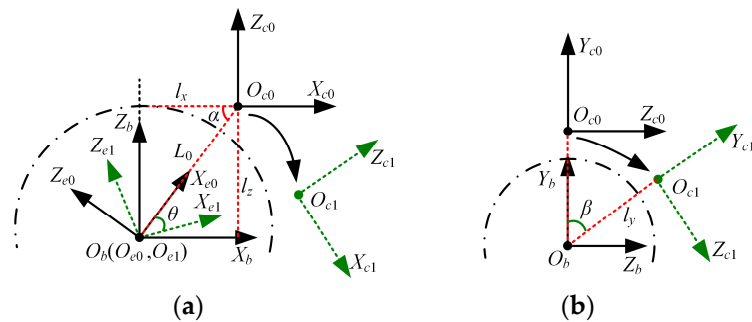
$$T_{c0}^{c1} = T_{e1}^{c1} T_{e0}^{e1} T_{c0}^{e0}, \tag{7}$$

where the expressions of T_{c0}^{c1}, T_{e1}^{c1}, T_{e0}^{e1}, and T_{c0}^{e0} are in Equation (8).

$$T_{c0}^{c1} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}, T_{e1}^{c1} = \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha & L_0\sin\alpha \\ 0 & 1 & 0 & 0 \\ \sin\alpha & 0 & \cos\alpha & -L_0\cos\alpha \\ 0 & 0 & 0 & 1 \end{bmatrix},$$
$$T_{e0}^{e1} = \begin{bmatrix} \cos\theta & 0 & -\sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ \sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \text{ and } T_{c0}^{e0} = \begin{bmatrix} \cos\alpha & 0 & \sin\alpha & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\alpha & 0 & \cos\alpha & L_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{8}$$

Figure 6. Calibration principle of the translation parameters. (**a**) Translation parameters in X_b and Z_b directions, (**b**) translation parameters in Y_b direction.

Equation (9) can be obtained by joining Equations (7) and (8).

$$\cos\alpha = \left(M^2 + 1\right)^{-\frac{1}{2}}, \sin\alpha = \sqrt{\frac{M^2}{M^2 + 1}}, \text{ and } L_0 = \frac{t_1}{\sin\alpha - \cos\alpha\sin\theta - \sin\alpha\cos\theta}, \tag{9}$$

where $M = \frac{-t_3\sin\theta - t_1\cos\theta + t_1}{-t_1\sin\theta + t_3\cos\theta - t_3}$. T_{c0}^{c1} can be solved using Equation (4), and θ is the known rotation angle. Therefore, l_x and l_z can be calculated using Equation (10).

$$l_x = L_0\cos\alpha, \text{ and } l_z = L_0\sin\alpha. \tag{10}$$

According to the above principle, partial translation parameters of the left and right cameras can be solved. The pose adjustment module L and the pose adjustment module R are adjusted in the left–right and front–back directions according to the corresponding negative calculated translation parameters. Upon completion of the adjustments, the remaining translation parameters in the vertical direction continue to be calibrated, which is similar to the principle of calibrating l_x and l_z.

We also take the left camera as an example. Based on the transformation relationship between the coordinate frame $O_{c0}X_{c0}Y_{c0}Z_{c0}$ and the coordinate frame $O_{c1}X_{c1}Y_{c1}Z_{c1}$ in Figure 6b, Equation (11) can be obtained.

$$T_{c0}^{c1} = T_b^{c1}T_{c0}^b = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\beta & \sin\beta & -l_y \\ 0 & -\sin\beta & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & l_y \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\beta & \sin\beta & \cos\beta l_y - l_y \\ 0 & -\sin\beta & \cos\beta & -\sin\beta l_y \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{11}$$

Therefore, the translation parameter l_y in the Y_b direction is

$$l_y = \frac{-t_3}{\sin\beta}, \tag{12}$$

where t_3 is the translation component in the Z_{c1} direction of $\boldsymbol{T}_{c0}^{c1}$, which can be solved using Equation (4). The translation parameters of the left and right cameras in the vertical direction are solved. Once the camera pose is adjusted through the adjustment module, it reaches the state where the optical center coincides with the rotation center.

4.2. Servo Motor Zero-Position Calibration and Adjustment

The calibration and adjustment of the zero-position of the servo motor refers to calculating the angle at which the zero-position of the servo motor rotates to the initial position, and then resetting the zero-position of the servo motor. This can ensure that the BBCPS returns to the initial position no matter what movement it performs.

Considering that the optical center coincides with the rotation center, the calibration principle of the angle for the left and upper motors is shown in Figure 7. The projection point of the ring center of the left calibration object in the left camera is $\boldsymbol{p}_0$ when the servo motors are at their zero-position. The projection point of the ring center of the left calibration object in the left camera is $\boldsymbol{p}_c$ when the servo motors are in the initial position of the BBCPS. The projection point of the ring center of the left calibration object in the left camera is $\boldsymbol{p}_t$ when the angle of the upper servo motor is γ_y. The horizontal difference between $\boldsymbol{p}_c$ and $\boldsymbol{p}_0$ is represented by Δx, and the vertical difference is Δy. According to the equal vertex angle theorem, the angle of the left servo motor γ_x and the angle of the upper servo motor γ_y can be calculated using Equation (13).

$$\gamma_x = \arctan\left|\frac{\Delta x}{f}\right|, \text{ and } \gamma_y = \arctan\left|\frac{\Delta y}{f}\right|, \tag{13}$$

where f is the focal length of the camera. In addition, the rotation direction of the left servo motor is defined as positive rotation when Δx is positive. The definition of the direction of the upper servo motor is the same as the left servo motor.
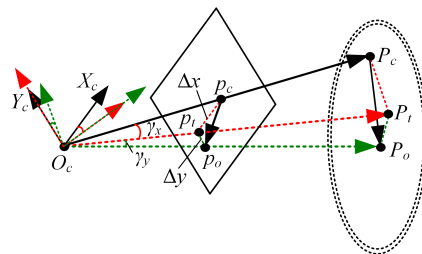

Figure 7. Calibration principle of the angle.

However, Equation (13) will fail when γ_x or γ_y is greater than a certain angle. When the servo motors are in their zero-position, the ring center may not be detected in the image because the ring center is beyond the field of view (FOV) of the left camera. Considering the above situation, we developed the procedure of the zero-position of the servo motor calibration and adjustment as shown in Figure 8. To prevent multiple circles from being detected, the background of the ring in the image should be kept simple, such as a pure-color wall. First, $\boldsymbol{p}_c$ is determined using the Hough circle detection algorithm [46] and recorded as the target point. Next, the left and upper servo motors are returned to their zero-position. To ensure the ring in the FOV, the left servo motor is required to rotate positively n and the upper servo motor needs to rotate positively m. The values of n and m are from $0°$ to $360°$. Subsequently, $\boldsymbol{p}_0$ can be detected using the Hough circle detection algorithm and the angles γ_x and γ_y can be calculated. The angles at which the left and upper servo motors rotate to the initial position are $n + \gamma_x$ and $m + \gamma_y$, respectively. Finally, the zero-positions of the left and upper servo motors are reset. The zero-position of the right servo motor can also be reset according to the above procedure.
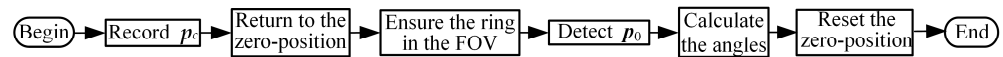
Figure 8. Procedure of the zero-position of the servo motor calibration and adjustment.

5. Control Method

In this section, the gaze mechanism of human eyes is imitated from the perspective of control. First, in our study, the motion-control method of a single servo motor based on a PID controller was introduced. Then, we developed a binocular interest-point extraction method based on frequency-tuned and template-matching algorithms. Furthermore, we proposed a binocular cooperative motion strategy to move the interest point to the principal point. Finally, real experiments were conducted to verify the effectiveness of the control method.

5.1. Motion Control of a Single Servo Motor

A PID controller is widely used in servo motor control because of its relative simplicity, easy adjustment, and fair performance [47]. The control principle is shown in Equation (14).

$$u(t) = K_P \cdot e(t) + K_i \int_0^t e(t)dt + K_d \frac{de(t)}{dt}, \tag{14}$$

where K_p is the proportional gain, K_i is the integral gain, K_d is the differential gain, t is the time, and $e(t)$ is the error.

By tuning K_p, K_i, and K_d, the motion of the servo motor can better follow the expectation, which is necessary for the control of the BBCPS. The requirement for parameter tuning is that the servo motor can achieve fast and stable motion and keep its motion error at about 0.1%. In response to this requirement, we experimentally determined the optimal K_p, K_i and K_d (the specific tuning process is described in Section 6.2.1).

5.2. Binocular Interest-Point Extraction Method

The frequency-tuned salient point detection algorithm is a classical method that analyzes an image from a frequency perspective to identify salient points in the image [34]. In this algorithm, the image needs to be Gaussian smoothed. Then, according to Equation (15), the saliency value of each pixel is calculated, and the pixel with the largest saliency value is the salient point.

$$S(p) = \left\| I_\mu - I(p) \right\|_2, \tag{15}$$

where I_μ represents the average feature of the image in Lab color space, $I(p)$ denotes the feature of the pixel point p in Lab color space, and $S(p)$ refers to the saliency value of the pixel point p. However, this algorithm can only determine the salient point in a single image. What we need to extract are interest points in binocular images.

When humans perceive a scene, one eye plays a leading role [48]. Inspired by this, we used the left camera as the leading eye, and proposed a binocular interest-point extraction method based on frequency-tuned and template-matching algorithms [49]. The flow of this method is shown in Algorithm 1.

The detailed description of Algorithm 1 is as follows. After inputting the left camera image I_l, the right camera image I_r, the image width W, and the template image width w_t, the frequency-tuned salient point detection algorithm is first used. The interest point $p_l = [x_l, y_l]$ in the image I_l is obtained. With the point p_l as the center, the template image I_t with a size of $w_t \times w_t$ is determined. A template-matching algorithm is used to match the corresponding interest point in the image I_r. Since the vertical visual field angles of the left and right cameras are consistent in the calibrated BBCPS, a matching algorithm through local search is performed to improve the speed. The starting location of the sliding window $[x_{smin}, y_{smin}]$ in the image I_r is defined as $[0, y_l - 2w_t]$, the ending location $[x_{smax}, y_{smax}]$ is $[W - w_t, y_l + 2w_t]$, and the sliding step d is 1. The similarity S between the sliding window and the template image at each position in the traversal interval is calculated using the

mean square error. We find the maximum S and record the corresponding location of the sliding window $[x_s, y_s]$. Finally, the interest point p_r in the right camera image is obtained.

Algorithm 1: Binocular interest-point extraction method based on frequency-tuned and template-matching algorithms.

Input: I_l, I_r, W pixels, w_t pixels
Output: p_l, p_r
Obtain p_l in I_l using the frequency-tuned algorithm [34]
Extract a template image I_t, with a size of $w_t \times w_t$, centered on p_l
for $y_s \leftarrow y_{smin}$ to y_{smax} step d **do**
  for $x_s \leftarrow x_{smin}$ to x_{smax} step d **do**
    Compute S and record $[S, x_s, y_s]$
  end for
end for
Find the maximum S and record the corresponding x_s and y_s
Obtain $p_r = [x_s + 0.5w_t, y_s + 0.5w_t]$

5.3. Binocular Cooperative Motion Strategy

To imitate the movement mechanism of the gaze function, we developed a binocular cooperative motion strategy. The implementation process of the strategy is shown in Figure 9. First, we used Algorithm 1 to extract the binocular interest point. Afterward, the rotation angle of the left servo motor γ_l, the rotation angle of the right servo motor γ_r, and the rotation angle of the upper servo motor γ_y were calculated. The principle of calculating the rotation angles of the three servo motors is shown in Figure 10.
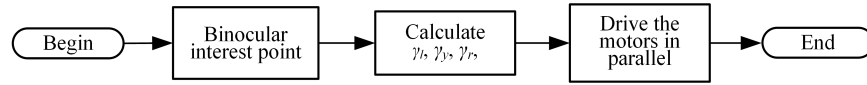

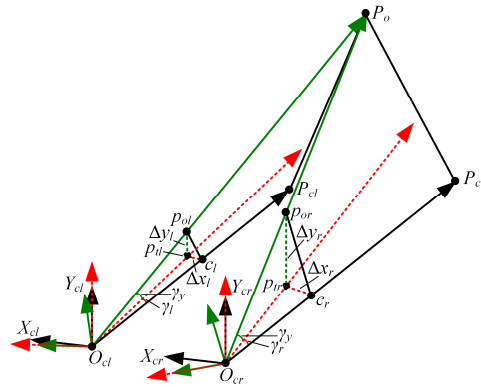**Figure 9.** Bionic binocular cooperative motion strategy.

Figure 10. The rotation angles of the three servo motors.

In this paper, we use $O_{cl}X_{cl}Y_{cl}Z_{cl}$ to define the left camera coordinate frame and $O_{cr}X_{cr}Y_{cr}Z_{cr}$ to describe the right camera coordinate frame. The points O_{cl} and O_{cr} are located at the left and right rotation centers. The 2D points p_{ol} and p_{or} are the projection points of the 3D interest point P_o in the left and right images, respectively. P_{cl} and P_{cr} are 3D points located on the optical axes of the left and right cameras. Their corresponding 2D points are the principal point of the left camera c_l and the principal point of the right camera c_r. The difference between c_r and p_{or} is denoted by Δx_r and Δy_r, and the difference between c_l and p_{ol} is described by Δx_l and Δy_l. The rotation angles of the three servo motors can be obtained via Equation (16). The rotation direction is specified in Section 4.2.

$$\gamma_l = \arctan\left|\frac{\Delta x_l}{f}\right|, \gamma_r = \arctan\left|\frac{\Delta x_r}{f}\right|, \text{ and } \gamma_y = \arctan\left|\frac{\Delta y_l + \Delta y_r}{2f}\right|. \tag{16}$$

Finally, the calculated rotation angles are sent to the three servo motors at the same time to realize the cooperative motion of the left and right cameras.

6. Experiments

In this part, we summarize the simulated and real experiments that were conducted. The simulated experiment was used to verify the superiority of the BBCPD assembled according to the principle of symmetrical distribution around the center compared to a device assembled according to the classical low center of gravity principle. Real experiments were then performed to test the control performance of BBCPS.
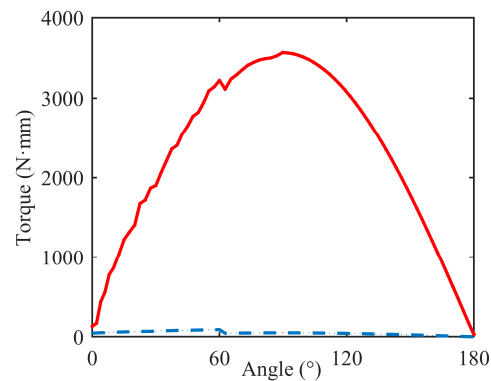
6.1. Simulation

The simulated experiment was performed to verify that the torque required by the upper servo motor in the BBCPD was less compared with that in the device shown in Figure 11. For this purpose, we used the SolidWorks Motion module to complete the simulation. We imported the 3D models in Figures 3 and 11 into the SolidWorks Motion module and entered the mass of each component. Then, we set the simulation parameters. The speed of the upper servo motor was $\frac{\pi}{3}$ rad·s^{-1}, and the acceleration time was 0.5 s. Generalized-alpha stability time-step integration for flexible figures (GSTIFF) [50] is an integral method of variable order and variable step size. This was selected as a type of integrator, suitable for various motion analysis problems. The initial integrator step was 1×10^{-4}, the minimum integrator step was 1×10^{-7}, and the maximum integrator step was 1×10^{-3}. The maximum number of iterations was 25. Afterward, the simulation was carried out. When the plane of the transmission frame was perpendicular to the plane of the base, the rotation angle of the upper servo motor was 0. The simulation results are shown in Figure 12.
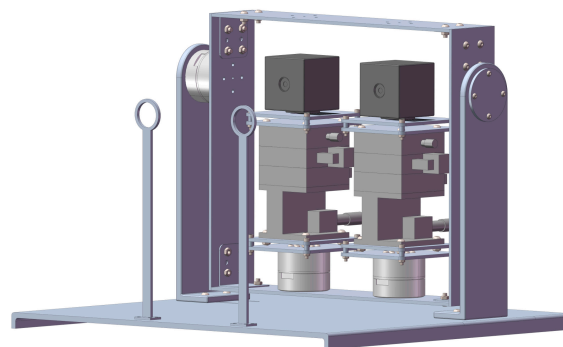
Figure 11. A device assembled according to the classical low center of gravity principle.

As the angle increases, the torques in the device and the BBCPD increase. When the angle was 90°, the torques reached their maximum values. The maximum torques in the BBCPD and the device were 49.5 N·m and 3565.8 N·m, respectively. This showed that the assembly principle of symmetrical distribution around the center can effectively reduce the maximum torque by 98.6%. When the angle was 0°, the torques in the BBCPD and the device were 46.9 N·m and 129.8 N·m, respectively. According to Equation (3), we find that the smaller the torque at the angle of 0°, the smaller the rotational inertia. Therefore, the inertia effect on the braking of the BBCPD is lower than that in the device. In addition, the torque could be effectively reduced by more than 97% in the BBCPD at any angle, which showed that the BBCPD greatly reduces energy consumption.

Figure 12. Simulation results, where the blue line represents the BBCPD and the red line represents the device shown in Figure 11.

6.2. Real Experiments

Next, we designed a real experiment to validate the control method of the BBCPS. The control method requires the BBCPD to complete initial position calibration. The initial position calibration method requires a known and accurate rotation angle of the servo motor. Therefore, the real experiment was divided into four parts. The first part of the experiment involved tuning the parameters of the PID controller to ensure that the servo motor could move stably and accurately. The second part of the experiment focused on verifying the effectiveness of the initial position calibration method, which lay a foundation for the subsequent experiments. In the third part of the experiments, we aimed to evaluate the accuracy of the proposed binocular interest-point extraction method, which can affect the control performance of the BBCPS. The accuracy of the binocular interest-point extraction method is limited by the accuracy of the template-matching algorithm. Thus, an experiment of template matching for images with viewpoint change was implemented. In the last part of the experiment, we verified the effectiveness of the binocular cooperative perception strategy and analyzed the control performance of the BBCPS.

To conduct these experiments, we constructed a real BBCPD, as shown in Figure 13, and corrected the lens distortions of the left and right cameras. The resolutions of the left and right cameras were 640 pixels × 480 pixels, their focal lengths were 532 pixels, and their principal points were [319 pixels, 242 pixels]. The three servo motors used were HT-S-4315, employing the RS485 communication mode. The rotation adjustment accuracy of the pose adjustment module R was 0.01°, and the translation adjustment accuracy was 0.01 mm. The configuration of the pose adjustment module L was the same as that of the pose adjustment module R.
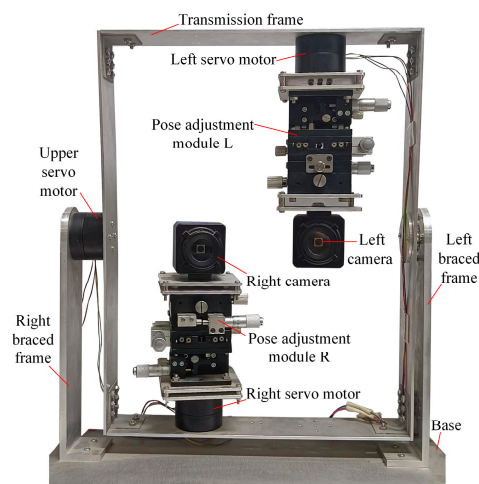
Figure 13. Real BBCPD.

6.2.1. PID Controller Parameter Tuning

In this part of our work, we used the approach of controlling variables to tune K_p, K_i, and K_d one by one to achieve fast and stable motion of the servo motor and maintain its motion error at about 0.1%. Taking the accuracy motion control of the servo motor to achieve $20° \pm 0.022°$ as an example, during the tuning process, we judged whether K_p, K_i, and K_d met the control requirement according to the motion steady-state response curve shown in Figure 14a.
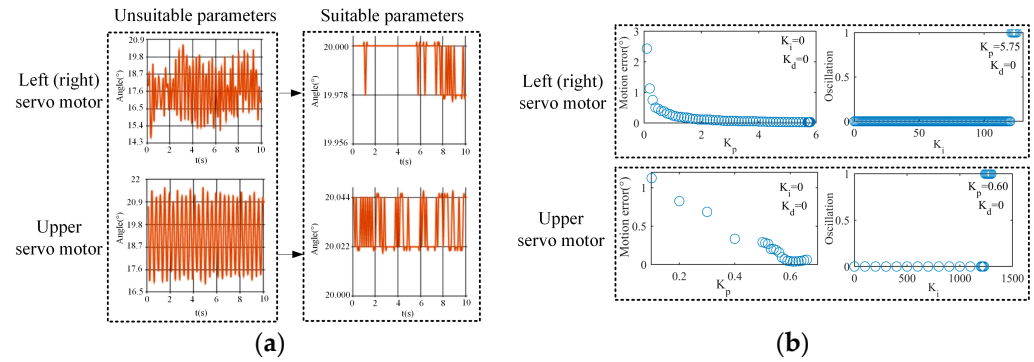
Figure 14. Parameter tuning. (**a**) Steady-state response curve of the motion, (**b**) process of tuning K_p, K_i, and K_d.

The parameter tuning process is shown in Figure 14b. First, we tuned K_p while setting K_i and K_d to zero. The increase in K_p can reduce the servo motor's motion error, but excessively large K_p may lead to overshooting, causing an increase in the motor's motion error. Therefore, we gradually increased K_p from 0 until overshooting occurred, and then we determined K_p. Subsequently, while keeping K_p and K_d constant, we tuned K_i. The increase in K_i can improve the response speed of the servo motor, but excessive K_i may introduce oscillations. Thus, K_i was gradually increased from 0 until oscillations occurred, and the maximum K_i without oscillation was taken. During this process, when the fluctuation range of the motion steady-state response curve exceeded $0.022°$, we regarded that oscillations occurred and were recorded as 1. K_d is used to suppress overshooting and oscillation, but an excessively large K_d can slow the response speed. Since K_p did not cause overshooting, K_d could be as small as possible to ensure the response speed. At that time, the parameter tuning was finished. For the right and left servo motors, K_p was 5.75, K_i was 113, and K_d was 0.5. For the upper servo motor, K_p was 0.60, K_i was 1240, and K_d was 0.4.

6.2.2. Initial Position Calibration of the BBCPD

The aim of this part of our work was to verify the effectiveness of the initial position calibration method for the BBCPD in Figure 13. First, the camera pose calibration and adjustment method was verified. According to the principle outlined in Section 4.1, we used an 8×11 checkerboard, shown in Figure 15, to calibrate the rotation parameters θ_x, θ_y, and θ_z. The calibration results are shown in Table 1. According to the calibrated parameters, the pose adjustment module R and the pose adjustment module L were adjusted to ensure that the photosensitive planes of the right and left two cameras were parallel. On this basis, we realized the calibration of the translation parameters l_x, l_y, and l_z. The calibration results are shown in Table 1. Adjustments to the pose adjustment modules were then performed according to the calibration results to ensure that the optical centers of the two cameras were located at the rotation centers. The adjustments of the two camera poses were completed.
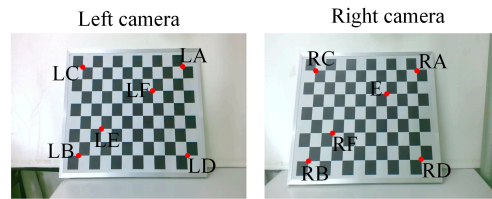
Figure 15. Selected corners.

Table 1. Camera pose calibration results.

Parameter	θ_x (°)	θ_y (°)	θ_z (°)	l_x (mm)	l_y (mm)	l_z (mm)
Left camera	−3.12	0.89	2.33	15.11	−20.42	16.27
Right camera	−1.56	1.45	−3.42	−9.07	−13.64	15.57

Since the adjustment of the camera rotation pose is the basis for the calibration of the translation parameters, it is only necessary to verify the effectiveness of the calibration and adjustment of the camera translation parameters. We used the checkerboard to perform six groups of validation experiments for each of the left and right cameras. As shown in Figure 15, images of the checkerboard were taken by the left and right cameras, respectively. We selected three pairs of corners in the image taken by the left camera (LA-LB, LC-LD, LE-LF) and three pairs of corners in the image taken by the right camera (RA-RB, RC-RD, RE-RF) as the experimental objects. The first group of experiments for the left camera was completed by using the camera to capture an image of the checkerboard after rotating 3° around the horizontal axis and 2° around the vertical axis. Similarly, the first group of experiments for the right camera was completed by using the camera to capture an image of the checkerboard after rotating 3° around the horizontal axis and 2° around the vertical axis. In the second to sixth groups of experiments for the left camera, rotation angles around the horizontal axis were {5°, 8°, 12°, 15°, 20°} and around the vertical axis were {3°, 5°, 9°, 13°, 16°}. For the second to sixth groups of experiments with the right camera, rotation angles around the horizontal axis were {5°, 8°, 12°, 15°, 20°} and around the vertical axis were {3°, 5°, 9°, 13°, 16°}.

E_t is defined as an evaluation index. The specific expression for E_t is presented in Equation (17).

$$E_t = \left| (\varphi_{i0} - \varphi_{j0}) - (\varphi_{i1} - \varphi_{j1}) \right|, \tag{17}$$

where φ_{i0} represents the angle between the light passing through a corner i and the optical axis before the camera rotates, φ_{i1} is the angle between the light passing through a corner i and the optical axis after the camera rotates, φ_{j0} refers to the angle between the light passing through corner j and the optical axis before the camera rotates, and φ_{j1} is the angle between the light passing through corner j and the optical axis after the camera rotates. If the optical center of the camera coincides with the rotation center, E_t is equal to 0 regardless of the degree of camera rotation. Through the Harris corner extraction algorithm, E_t in each group was calculated. The experimental results are shown in Figure 16.
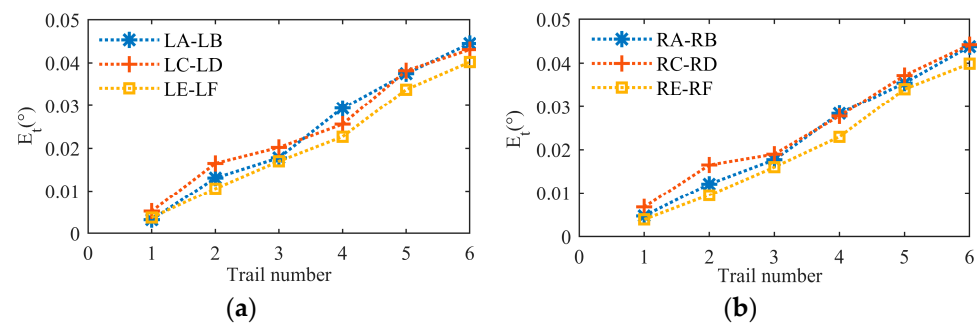

Figure 16. Calibration errors after camera pose adjustment. (**a**) Left camera, (**b**) right camera.

E_t was not equal to 0. This indicated that the camera center did not completely coincide with the rotation center. Further observation revealed that the maximum E_t did not exceed 0.05°. When converting E_t into pixel information based on the known camera focal length, the pixel error generated by an error of 0.05° was less than 1. Considering the corner detection error, the servo motor motion error, the adjustment error of the pose adjustment module, and other factors, we surmised that an error below 1 pixel can imply the effectiveness of the camera pose calibration and adjustment method. Additionally, by combining the rotation angle of the camera in each group of experiments, we observed that as the rotation angle increased, E_t also increased. The reason for this is that the optical centers of the left and right cameras do not perfectly coincide with the rotation centers, and the selected corner points are not on the same depth plane. Corners at different depths generate different motion disparities after the camera moves. With an increase in the rotation angle, the difference in motion disparity between two corners at different depths increases, leading to the increase in E_t.

Next, the calibration and adjustment method for the servo motor zero-position was verified. According to the calibration principle of Section 4.2, the process of the calibration and adjustment for the three servo motors is shown in Figure 17. First, the pixel coordinates of the ring center p_c were determined through Hough circle detection, when the BBCPD was in the ideal initial position. The three servo motors returned to their zero-positions. We detected the pixel coordinates of the ring center p_0. According to Equation (13), the calibrated angles for the left, upper, and right servo motors were determined as 26.2°, 8.5°, and 19.6°, respectively. Based on the calibrated angles, the zero-positions of the three servo motors were reset.
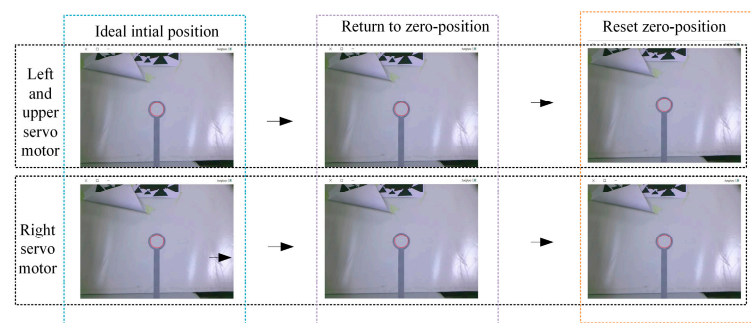
Figure 17. The process of calibration and adjustment of the servo motor zero-position.

To validate the adjustment effectiveness of the servo motor zero-position, we conducted six groups of experiments on the three servo motors following the "rotating-zeroing" procedure. In each group of experiments, the rotations of the three servo motors were the same. The rotation angles of the three servo motors in the six groups of experiments were {3°, 5°, 6°, 15°, 10°, 13°}. We defined the Euclidean distance between p_c and the pixel coordinates of the ring center detected from the image taken by the left camera when the servo motors returned to the zero-positions as E_l. Similarly, for the right camera, this was denoted as E_r. The experimental results are listed in Table 2. After multiple motions of the servo motors, the average E_r and the average E_l were both less than 1 pixel. The adjustment of the servo motor zero-position was feasible, with E_r and E_l having little effect on the subsequent control performance. Further analysis was not conducted.

Table 2. Calibration errors after servo motor zero-position adjustment.

Trail Number	1	2	3	4	5	6
E_l (pixel)	1.0	0.0	0.0	1.4	1.4	1.0
E_r (pixel)	0.0	1.0	1.4	1.0	1.0	1.4

6.2.3. Template Matching for Images with Viewpoint Change

To test the accuracy of our template-matching algorithm, six groups of experiments were carried out. As shown in Figure 18, experimental images were selected from the publicly available dataset HPatches [51], featuring images with various viewpoints. We used the FT algorithm to extract the pixel coordinates of an interest point in an image. Subsequently, a template with a size of 60 pixels $\times$ 60 pixels was used to match this interest point in another image, and the matched point was recorded as p_i. The dataset provides a mapping relationship between images, enabling the calculation of the position of this interest point in another image, which is considered the ground truth p_r. We define E_m as an evaluation index, as shown in Equation (18).
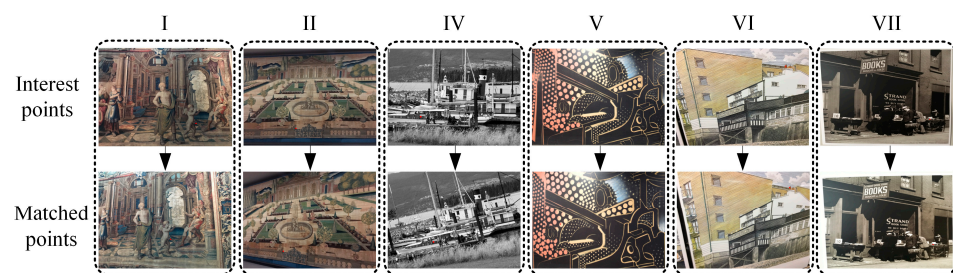
$$E_m = \|p_i - p_r\|_2. \tag{18}$$

Figure 18. Template-matching experiment.

The results of the experiments are presented in Table 3. Considering the existence of estimation errors in the mapping matrix, it is generally accepted that a condition for correct matching is E_m within 3 pixels. Except for the fifth group of experiments, the E_m values of all other experiments were less than 3 pixels. This indicates that our template-matching algorithm is suitable for matching images with different viewpoints. Combined with the images in Figure 18, we analyzed the fifth group of experiments. It can be observed that the images in the fifth group of experiments have larger variations in viewpoint compared to the other groups. This suggests that our template-matching algorithm may not be suitable for scenes with significant viewpoint changes, which is a direction for our future research. For other groups where images had smaller variations in viewpoint, the average E_m was less than 0.8 pixels. This indicates that our template-matching algorithm performs well when the scene viewpoint does not vary significantly, with an error of around 1 pixel.

Table 3. Template-matching results.

Trail Number	1	2	3	4	5	6
p_i (pixel)	(610, 594)	(340, 696)	(293, 496)	(325, 561)	(821, 210)	(825, 975)
p_r (pixel)	(608, 594)	(339, 696)	(293, 495)	(325, 561)	(822, 213)	(825, 975)
E_m (pixel)	2.0	1.0	1.0	0.0	3.2	0.0

6.2.4. Binocular Cooperative Perception

In this part of our work, we conducted six groups of experiments to verify our binocular cooperative perception method. The experimental process is shown in Figure 19. Taking the second group of experiments as an example, the left and right cameras took images at the initial position. Next, the pixel coordinates of the interest points in the left and right cameras were extracted using the binocular interest-point extraction method outlined in Section 5.2. In this process, the size of a template image was set to 30 pixels $\times$ 30 pixels. Finally, the motion angles of the three servo motors were calculated using Equation (16), and then the three servo motors were driven in parallel to set the gaze on the interest point. To avoid randomness, we conducted another five groups of experiments following the same steps.
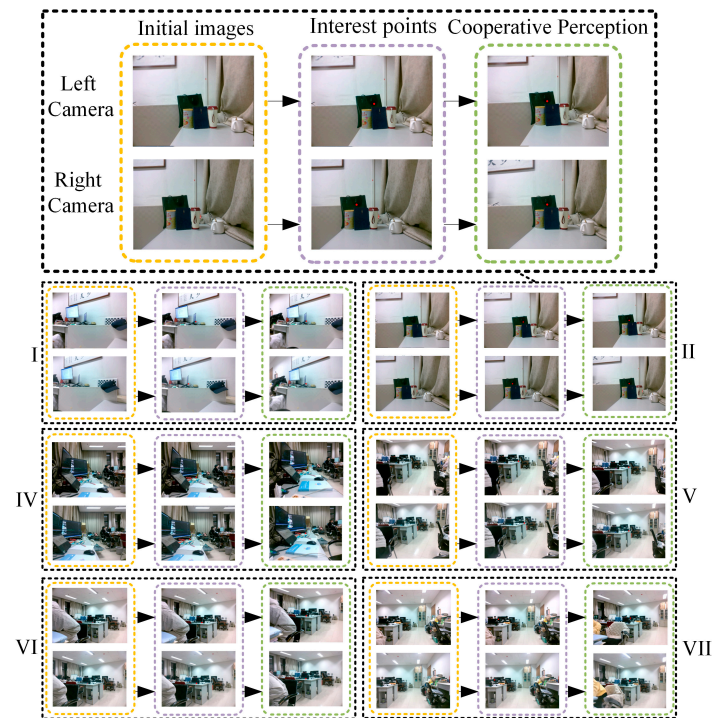
Figure 19. Binocular cooperative perception experiment.

The gaze error E_g is defined as an evaluation index, as expressed in Equation (19).

$$E_g = \left\| \boldsymbol{p}_{gl} - \boldsymbol{c}_l \right\|_2 + \left\| \boldsymbol{p}_{gr} - \boldsymbol{c}_r \right\|_2, \tag{19}$$

where $\boldsymbol{p}_{gl}$ represents the pixel coordinates of the interest point in the left image after gazing, $\boldsymbol{p}_{gr}$ comprises the pixel coordinates of the interest point in the left image after gazing, and $\boldsymbol{c}_l$ and $\boldsymbol{c}_r$ are the principal points of the left and right cameras. The experimental results are shown in Figure 20.

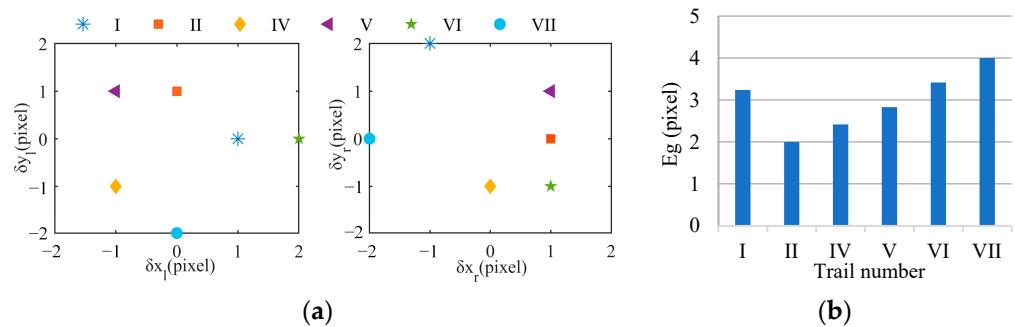**Figure 20.** Errors of binocular cooperative perception. (**a**) Error distribution, (**b**) gaze error.

Figure 20a illustrates the distribution of the gaze error. The symbols δx_r and δy_r are the difference between $\boldsymbol{c}_r$ and $\boldsymbol{p}_{gr}$, and the meanings of δx_l and δy_l are the same as those of δx_r and δy_r. Figure 20b shows the E_g of each experiment. We found that the absolute values of δx_l, δy_l, δx_r, and δy_r in each group of experiments were less than 2 pixels, and the average E_g was less than 3 pixels. This indicates that the perceived interest point basically coincides with the principal point, and it confirms the effectiveness of our control method. The error is attributed to the static error of the servo motor, the template-matching error, and the initial position calibration error. The static error of the servo motor can lead the actual motion angles of the servo motors to deviate from the calculated γ_l, γ_r, and γ_y. The template-matching error results in the matched interest point in the right image deviating

from the true corresponding interest point. This means that the calculated γ_y is not the angle needed in order for the interest point completely coincide with the principal point. The initial position calibration error introduces the interference of the depth information of the interest point. The difference motion disparity results in a difference between the calculated angle and the angle needed for the interest point to completely coincide with the principal point.

Further observation found that δx_l, δy_l, δx_r, and δy_r fluctuated in the six groups of experiments. The reason is that, in different groups, the static error of the servo motor varies, the template-matching error differs, and the influence degree of the initial position calibration error is also different. In each group of experiments, δx_l, δy_l, δx_r, and δy_r also exhibited fluctuations. These were because the static errors of the left and right servo motors may differ and the camera pose adjustment errors for the left and right cameras may also vary. We also noted that E_g was the smallest in the second group of experiments, at 2 pixels, while E_g was the largest in the sixth group of experiments, at 4 pixels. Our analysis indicates that the viewpoint change in the images affects E_g. With an increase in the viewpoint change, the rotation angle increases, and the static error of the servo motor also increases. This can lead to a larger difference between the calculated angle and the actual motion angle, potentially resulting in an increase in E_g. Moreover, an increase in the viewpoint change enlarges the error of the initial position calibration, and then the interest point generates a larger motion disparity. This can lead to a larger difference between the calculated angle and the angle needed for the interest point to completely coincide with the principal point, potentially further increasing E_g.

7. Conclusions

In this study, motivated by the eye gaze mechanism, we designed the flexible BBCPD. The device was assembled according to the principle of symmetrical distribution around the center based on dynamic analysis. The innovative principle offers distinct advantages by enhancing braking performance and reducing energy consumption in comparison to the classic symmetrical low center of gravity principle (as shown in Figure 11). A simulation was conducted to verify the advantages. The results showed that the innovative principle could reduce the torque of the upper servo motor by more than 97%, which leads to a reduction in energy consumption of the BBCPD. The results also demonstrated that the principle could lead the BBCPD to have smaller rotational inertia of the load of the upper servo motor, thus enhancing the braking performance of the BBCPD.

Furthermore, we developed an initial position calibration technique for the BBCPD. Based on the calibration results, the BBCPD, after adjusting the pose adjustment modules and resetting the zero-positions of the servo motors, meets the requirement of the control method. Subsequently, the control method was proposed, where a binocular interest-point extraction method based on frequency-tuned and template-matching algorithms was applied to detect the interest points. Then, we crafted a binocular cooperative motion-control strategy for how servo motors could coordinate their movements and thus set the gaze upon an interest point. Last, real experiments were conducted, and the results showed that the control method of the BBCPS could achieve a gaze error within 3 pixels.

The proposed BBCPS can advance the development of humanoid intelligent perception, with application prospects in fields such as intelligent manufacturing [52,53], human–robot interaction [5], and autonomous driving [54]. However, the gaze accuracy of the BBCPS may constrain its further development. In the future, we aim to reduce gaze errors by optimizing our control algorithm. For instance, by referring to previous research on image matching under viewpoint changes [55,56], we plan to improve the matching algorithm in order to enhance the precision of the binocular interest-point extraction algorithm.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available to protect the privacy of the subjects.

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

References

1. Cheng, Y.H.; Zhang, X.C.; Lu, F.; Sato, Y. Gaze Estimation by Exploring Two-Eye Asymmetry. *IEEE Trans. Image Process.* **2020**, *29*, 5259–5272. [CrossRef]
2. Grotz, M.; Habra, T.; Ronsse, R.; Asfour, T. Autonomous View Selection and Gaze Stabilization for Humanoid Robots. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 24–28 September 2017; pp. 1427–1434.
3. Pateromichelakis, N.; Mazel, A.; Hache, M.A.; Koumpogiannis, T.; Gelin, R.; Maisonnier, B.; Berthoz, A. Head-eyes system and gaze analysis of the humanoid robot Romeo. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Chicago, IL, USA, 14–18 September 2014; pp. 1374–1379.
4. Belkaid, M.; Kompatsiari, K.; De Tommaso, D.; Zablith, I.; Wykowska, A. Mutual gaze with a robot affects human neural activity and delays decision-making processes. *Sci. Robot.* **2021**, *6*, 5044. [CrossRef]
5. Saran, A.; Majumdar, S.; Short, E.S.; Thomaz, A.; Niekum, S. Human Gaze Following for Human-Robot Interaction. In Proceedings of the 25th IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2018; pp. 8615–8621.
6. Qiu, Y.N.; Busso, C.; Misu, T.; Akash, K. Incorporating Gaze Behavior Using Joint Embedding with Scene Context for Driver Takeover Detection. In Proceedings of the 47th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 22–27 May 2022; pp. 4633–4637.
7. Yoshida, S.; Yoshikawa, M.; Sangu, S. Autonomous calibration for gaze detection using Bayesian estimation and canonical correlation analysis. In Proceedings of the Conference on Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR) III, San Francisco, CA, USA, 23–25 January 2022; p. 1193103.
8. Marg, E. Development of electro-oculography; standing potential of the eye in registration of eye movement. *AMA Arch. Ophthalmol.* **1951**, *2*, 169–185. [CrossRef]
9. Carter, B.T.; Luke, S.G. Best practices in eye tracking research. *Int. J. Psychophysiol.* **2020**, *155*, 49–62. [CrossRef]
10. Tatara, S.; Toda, H.; Maeda, F.; Handa, T. Development of a New Eye Movement Measurement Device Using Eye-Tracking Analysis Technology. *Appl. Sci.* **2023**, *13*, 5968. [CrossRef]
11. Henderson, J.M.; Luke, S.G. Stable Individual Differences in Saccadic Eye Movements During Reading, Pseudoreading, Scene Viewing, and Scene Search. *J. Exp. Psychol.-Hum. Percept. Perform.* **2014**, *40*, 1390–1400. [CrossRef]
12. Sargezeh, B.A.; Tavakoli, N.; Daliri, M.R. Gender-based eye movement differences in passive indoor picture viewing: An eye-tracking study. *Physiol. Behav.* **2019**, *206*, 43–50. [CrossRef]
13. Chen, Z.L.; Chang, K.M. Cultural Influences on Saccadic Patterns in Facial Perception: A Comparative Study of American and Japanese Real and Animated Faces. *Appl. Sci.* **2023**, *13*, 11018. [CrossRef]
14. Goliskina, V.; Ceple, I.; Kassaliete, E.; Serpa, E.; Truksa, R.; Svede, A.; Krauze, L.; Fomins, S.; Ikaunieks, G.; Krumina, G. The Effect of Stimulus Contrast and Spatial Position on Saccadic Eye Movement Parameters. *Vision* **2023**, *7*, 68. [CrossRef]
15. Bang, Y.B.; Paik, J.K.; Shin, B.H.; Lee, C. A Three-Degree-of-Freedom Anthropomorphic Oculomotor Simulator. *Int. J. Control Autom. Syst.* **2006**, *4*, 227–235. [CrossRef]
16. 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.* **2014**, *78*, 425–441. [CrossRef]
17. Xie, Y.H.; Liu, J.Y.; Li, H.Y.; Han, C.; Xie, S.R.; Luo, J. Design and validation of robotic bionic eye with multiple flexible ropes parallel mechanism inspired by oculomotor law. *Mechatronics* **2021**, *80*, 102686. [CrossRef]
18. Rajendran, S.K.; Wei, Q.; Zhang, F.T. Two degree-of-freedom robotic eye: Design, modeling, and learning-based control in foveation and smooth pursuit. *Bioinspir. Biomim.* **2021**, *16*, 046022. [CrossRef]
19. Wang, X.Y.; Zhang, Y.; Fu, X.J.; Xiang, G.S. Design and Kinematic Analysis of a Novel Humanoid Robot Eye Using Pneumatic Artificial Muscles. *J. Bionic Eng.* **2008**, *5*, 264–270. [CrossRef]

20. Fan, D.; Chen, X.P.; Zhang, T.R.; Chen, X.; Liu, G.L.; Owais, H.M.; Kim, H.; Tian, Y.; Zhang, W.M.; Huang, Q. Design of Anthropomorphic Robot Bionic Eyes. In Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO), Macau, China, 5–8 December 2017; pp. 2050–2056.

21. Hofmann, J.; Domdei, L.; Jainta, S.; Harmening, W.M. Assessment of binocular fixational eye movements including cyclotorsion with split-field binocular scanning laser ophthalmoscopy. *J. Vision* **2022**, *22*, 1–13. [CrossRef]

22. 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]

23. Huang, C.J.; Gu, J.; Luo, J.; Li, H.Y.; Xie, S.R.; Liu, H.L. System Design and Study of Bionic Eye Based on Spherical Ultrasonic Motor Using Closed-loop Control. In Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO), Shenzhen, China, 12–14 December 2013; pp. 2685–2690.

24. Flores, E.; Fels, S. A Novel 4 DOF Eye-camera Positioning System for Androids. In Proceedings of the 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Kobe, Japan, 31 August–4 September 2015; pp. 621–627.

25. Blatnicky, M.; Dizo, J.; Sága, M.; Gerlici, J.; Kuba, E. Design of a Mechanical Part of an Automated Platform for Oblique Manipulation. *Appl. Sci.* **2020**, *10*, 8467. [CrossRef]

26. Meng, Q.; Xu, R.; Xie, Q.; Bostan·Mahmutjan; Li, S.; Yu, H. Bionic Design to Reduce Driving Power for a Portable Elbow Exoskeleton Based on Gravity-balancing Coupled Model. *J. Bionic Eng.* **2022**, *20*, 146–157. [CrossRef]

27. Ghosh, B.K.; Wijayasinghe, I.B. Dynamic Control of Human Eye on Head System. In Proceedings of the 29th Chinese Control Conference, Beijing, China, 29–31 July 2010; pp. 5514–5519.

28. Mao, X.B.; Chen, T.J. A biologically inspired model of robot gaze shift control. In Proceedings of the International Conference on Computers, Communications, Control and Automation (CCCA 2011), Hong Kong, China, 20–21 February 2011; pp. 185–189.

29. Oki, T.; Ghosh, B.K. Stabilization and Trajectory Tracking of Version and Vergence Eye Movements in Human Binocular Control. In Proceedings of the European Control Conference (ECC), Linz, Austria, 15–17 July 2015; pp. 1573–1580.

30. Muhammad, W.; Spratling, M.W. A Neural Model of Coordinated Head and Eye Movement Control. *J. Intell. Robot. Syst.* **2017**, *85*, 107–126. [CrossRef]

31. Wang, Q.B.; 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]

32. Rubies, E.; Palacín, J.; Clotet, E. Enhancing the Sense of Attention from an Assistance Mobile Robot by Improving Eye-Gaze Contact from Its Iconic Face Displayed on a Flat Screen. *Sensors* **2022**, *22*, 4282. [CrossRef] [PubMed]

33. Hou, X.D.; Zhang, L.Q. Saliency detection: A spectral residual approach. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007; pp. 1–8.

34. Achanta, R.; Hemami, S.; Estrada, F.; Süsstrunk, S. Frequency-tuned Salient Region Detection. In Proceedings of the IEEE-Computer-Society Conference on Computer Vision and Pattern Recognition Workshops, Miami Beach, FL, USA, 20–25 June 2009; pp. 1597–1604.

35. Gao, B.; Kou, Z.M.; Jing, Z.M. Stochastic Context-Aware Saliency Detection. In Proceedings of the 2011 International Conference on Computer and Management (CAMAN), Wuhan, China, 19–21 May 2011; pp. 1–5.

36. Yan, K.; Wang, X.Y.; Kim, J.M.; Zuo, W.M.; Feng, D.G. Deep Cognitive Gate: Resembling Human Cognition for Saliency Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 4776–4792. [CrossRef]

37. Li, Q. Saliency prediction based on multi-channel models of visual processing. *Mach. Vis. Appl.* **2023**, *34*, 47. [CrossRef]

38. Malladi, S.P.K.; Mukherjee, J.; Larabi, M.C.; Chaudhury, S. Towards explainable deep visual saliency models. *Comput. Vis. Image Underst.* **2023**, *235*, 103782. [CrossRef]

39. Zhu, Q.P.; Triesch, J.; Shi, B.E. Joint Learning of Binocularly Driven Saccades and Vergence by Active Efficient Coding. *Front. Neurorobotics* **2017**, *11*, 58. [CrossRef]

40. Iskander, J.; Hossny, M.; Nahavandi, S. A Review on Ocular Biomechanic Models for Assessing Visual Fatigue in Virtual Reality. *IEEE Access* **2018**, *6*, 19345–19361. [CrossRef]

41. Iskander, J.; Hossny, M.; Nahavandi, S.; del Porto, L. An ocular biomechanic model for dynamic simulation of different eye movements. *J. Biomech.* **2018**, *71*, 208–216. [CrossRef] [PubMed]

42. Shin, H.J.; Lee, S.J.; Oh, C.S.; Kang, H. Novel compact device for clinically measuring extraocular muscle (EOM) tension. *J. Biomech.* **2020**, *109*, 109955. [CrossRef]

43. Pallus, A.C.; Walton, M.M.G.; Mustari, M.J. Response of supraoculomotor area neurons during combined saccade-vergence movements. *J. Neurophysiol.* **2018**, *119*, 585–596. [CrossRef]

44. Zhang, Z.Y. A Flexible New Technique for Camera Calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1330–1334. [CrossRef]

45. Hou, Y.L.; Su, X.Y.; Chen, W.J. Axis alignment method in the rotating photogrammetric system. *Opt. Eng.* **2021**, *60*, 064105. [CrossRef]

46. Li, Q.; Wu, M.Y. An Improved Hough Transform for Circle Detection using Circular Inscribed Direct Triangle. In Proceedings of the 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Chengdu, China, 17–19 October 2020; pp. 203–207.

47. Zawawi, M.; Elamvazuthi, I.; Aziz, A.B.A.; Daud, S.A. Comparison of PID and Fuzzy Logic Controller for DC Servo Motor in the development of Lower Extremity Exoskeleton for Rehabilitation. In Proceedings of the IEEE 3rd International Symposium in Robotics and Manufacturing Automation (ROMA), Kuala Lumpur, Malaysia, 19–21 September 2017; pp. 1–6.

48. Momeni-Moghaddam, H.; McAlinden, C.; Azimi, A.; Sobhani, M.; Skiadaresi, E. Comparing accommodative function between the dominant and non-dominant eye. *Graefes Arch. Clin. Exp. Ophthalmol.* **2014**, *252*, 509–514. [CrossRef] [PubMed]

49. Yoo, J.; Hwang, S.S.; Kim, S.D.; Ki, M.S.; Cha, J. Scale-invariant template matching using histogram of dominant gradients. *Pattern Recognit.* **2014**, *47*, 3006–3018. [CrossRef]

50. Gear, C. Simultaneous Numerical Solution of Differential-Algebraic Equations. *IEEE Trans. Circuit Theor.* **1971**, *18*, 89–95. [CrossRef]

51. Balntas, V.; Lenc, K.; Vedaldi, A.; Mikolajczyk, K. HPatches: A Benchmark and Evaluation of Handcrafted and Learned Local Descriptors. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 3852–3861.

52. Kuric, I.; Kandera, M.; Klarák, J.; Ivanov, V.; Wiecek, D. Visual Product Inspection Based on Deep Learning Methods. In Proceedings of the Grabchenko's International Conference on Advanced Manufacturing Processes (InterPartner), Odessa, Ukraine, 10–13 September 2019; pp. 148–156.

53. Zhou, G.H.; Zhang, C.; Li, Z.; Ding, K.; Wang, C. Knowledge-driven digital twin manufacturing cell towards intelligent manufacturing. *Int. J. Prod. Res.* **2020**, *58*, 1034–1051. [CrossRef]

54. Chen, G.; Wang, F.; Li, W.J.; Hong, L.; Conradt, J.; Chen, J.N.; Zhang, Z.Y.; Lu, Y.W.; Knoll, A. NeuroIV: Neuromorphic Vision Meets Intelligent Vehicle Towards Safe Driving With a New Database and Baseline Evaluations. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 1171–1183. [CrossRef]

55. Božek, P.; Pivarčiová, E. Registration of holographic images based on the integral transformation. *Comput. Inform.* **2013**, *31*, 1369–1383.

56. Xia, X.H.; Xiang, H.M.; Cao, Y.S.; Ge, Z.K.; Jiang, Z.A. Feature Extraction and Matching of Humanoid-Eye Binocular Images Based on SUSAN-SIFT Algorithm. *Biomimetics* **2023**, *8*, 139. [CrossRef]