

Research Article

Symmetric Kullback-Leibler Metric Based Tracking Behaviors for Bioinspired Robotic Eyes

Hengli Liu, Jun Luo, Peng Wu, Shaorong Xie, and Hengyu Li

School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200072, China

Correspondence should be addressed to Hengyu Li; lihengyu@shu.edu.cn

Received 8 July 2015; Revised 5 October 2015; Accepted 21 October 2015

Academic Editor: Cecilia Laschi

Copyright © 2015 Hengli Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A symmetric Kullback-Leibler metric based tracking system, capable of tracking moving targets, is presented for a bionic spherical parallel mechanism to minimize a tracking error function to simulate smooth pursuit of human eyes. More specifically, we propose a real-time moving target tracking algorithm which utilizes *spatial histograms* taking into account symmetric Kullback-Leibler metric. In the proposed algorithm, the key spatial histograms are extracted and taken into particle filtering framework. Once the target is identified, an image-based control scheme is implemented to drive bionic spherical parallel mechanism such that the identified target is to be tracked at the center of the captured images. Meanwhile, the robot motion information is fed forward to develop an adaptive smooth tracking controller inspired by the *Vestibuloocular Reflex* mechanism. The proposed tracking system is designed to make the robot track dynamic objects when the robot travels through transmittable terrains, especially bumpy environment. To perform *bumpy-resist capability* under the condition of violent attitude variation when the robot works in the bumpy environment mentioned, experimental results demonstrate the effectiveness and robustness of our bioinspired tracking system using bionic spherical parallel mechanism inspired by head-eye coordination.

1. Introduction

Robot vision systems are crucial to recognize and acquire surrounding information for mobile robots. Target tracking, target recognition, surrounding perception, robotic localization, and attitude estimation are the most popular topics in robotics. And the target tracking function has emerged as a significant aspect for *Human Robot Interaction (HRI)*, camera *Motion-Disturbance Compensation (MDC)*, and tracking stabilization.

The robot motion information is commonly used to keep the camera stabilization and compensate small rotation or movements of the camera. These systems used inertial sensors and visual cues to compute the motion information of the camera. Jung and Sukhatme [1] developed a *Kanade-Lucas-Tomasi (KLT)* based motion tracking system for a moving target using a single camera on a mobile robot. Hwangbo et al. [2, 3] also developed a gyro-aided KLT feature tracking method that remained robust under fast camera-ego rotation conditions. Park et al. [4] proposed an Extended Kalman Filter (*EKF*) based motion data fusion scheme for

visual object tracking by autonomous vehicles. Jia et al. [5] also proposed a scheme of joint of visual features and the vehicle's inertial measurements for visual object identification and tracking. Hol et al. [6] used a multirate EKF by fusing measurements from inertial sensors (accelerometers and rate gyroscopes) and vision to estimate and predict position and orientation (pose) of a camera for robust real-time tracking.

Recently, biomimetic systems were extensively investigated by adopting the movement mechanics of human eye. The development of eyeball's neurophysiology provides a large amount of data and theory foundation for building up the controlling model of eye movement. Among the several types of eye movements, smooth tracking and gaze stabilization play a fundamental role. Lenz et al. [7] developed an adaptive gaze stabilization controller inspired by the *Vestibuloocular Reflex (VOR)*. It integrated inertial and visual information to drive the eyes in the opposite direction to head movement and thereby stabilized the image on the retina under dynamic changes. Shibata's biological oculomotor systems [8] used human-eye's VOR and Optokinetic Reflex

(OKR) to improve the gaze stabilization of vision system. A chameleon-inspired binocular “negative correlation” visual system (CIBNCVS) with neck [9] was designed to achieve swift and accurate positioning and tracking. Avni et al. [10] also presented a biologically motivated approach of tracking with independent cameras inspired by chameleon-like visual system. Law et al. [11] described biologically constrained architecture for developmental learning of eye-head gaze control on an iCub robot. Xie et al. [12] proposed a biomimetic control strategy of on-board pan-tilt-zoom camera to stabilize visual tracking from a helicopter based on physiological neural path of eye movement control. Vannucci et al. [13] established an adaptive model for robotic control able to perform visual pursuit with prediction of the target motion. Falotico et al. [14] employed “catch-up” saccade model to fixate the object of interest in case of moving targets in order to obtain a human-like tracking system. Compared with the classical control methods, the advantages of using a bionic controller make the robot easily adapted to transmittable terrains and track moving targets stably. Inspired by the excellent work, we tackle turbulence problem of tracking when the robots travel through bumpy terrains using a tracking system, that is, *bumpy-resist capability*.

Furthermore, with the development of anatomy of human eye, the movement mechanics of the human eye have aroused much interest in bionic engineering. Humanoid robot James [15, 16] was equipped with two artificial eyes, which can pan and tilt independently (totally 4 DOFs.). Thus, the iCub [17, 18] also had two artificial eyes with 3 DOFs, offering viewing and tracking motions. Wang et al. [19] devised a novel humanoid robot eye, which is driven by six pneumatic artificial muscles (PAMs) and rotates with 3 DOFs. Bioinspired actuators and mechanisms have been proposed to pan and tilt a camera with comparable characteristics as a human eye [20, 21]. Tendon-driven robot eye [22] was presented utilizing a mechanical base of the geometry of the eye and of its actuation system behind the implementation of Listing’s law. Gu et al. [23] presented an artificial eye implant with shape memory alloys (SMAs) driven by a small servomotor. A miniature artificial compound eye called the curved artificial compound eye (CurvACE) [24] was endowed using similar micromovements to those occurring in the fly’s compound eye.

Many bionic eyes have been presented as mentioned above. However, spherical parallel mechanism (SPM) has a compact structure, excellent dynamic performance, and high accuracy; in addition, a 3-DOF SPM is in line with the structural design of the bionic eye. 3-DOF SPMs attract decent amount of interest for this reason. A large number of these 3-DOF SPM bionic eyes have been proposed. An artificial eye [25, 26] for humanoid robots has been devised to be small in size and weight as well as to imitate the high dynamic movements of the human eye. The “Agile Eye” [27] is a high-performance parallel mechanism which has the capability of orienting a camera mounted end effector within a workspace larger than that of a human eye and with velocities and accelerations larger than those of the human eye. Bang et al. [28] design a 3-DOF anthropomorphic oculomotor system to match the human-like eye’s performance capabilities. Our

mechanism platform is inspired by these excellent works and plays a vital role in tracking dynamic objects.

Tracking a dynamic object when a robot performs its normal motion is common in application. To keep smoothly tracking moving objects, we develop a *bioinspired tracking system that is extensively used when the robot works in bumpy environment or with dynamic disturbance* in this paper. With active robot vision, an image-based feedback tracking system is presented for our bionic SPM to minimize tracking servoing, capable of tracking moving target when the robot moves across in bumpy environment. More specifically, we propose a real-time moving target tracking algorithm which utilizes spatial histograms and symmetric Kullback-Leibler (SKL) metric integrated in particle filtering framework to achieve automatic moving target tracking and gaze stabilization. In the proposed algorithm, the key spatial histograms are extracted and taken into particle filtering framework. An image-based feedback control scheme is implemented to drive bionic SPM such that the identified target is to be tracked at the center of the captured images. Meanwhile, the robot motion information is fed forward to develop an adaptive smooth tracking controller bioinspired by the VOR mechanism. To perform good specification, we test our vision stability system under the condition of violent attitude variation when the robot works in *bumpy environment*.

From a robotics point of view, our system is biologically inspired. While smooth tracking is employed to create a consistent perception of the surrounding world, the interaction with environment is also used to adjust the control model involved in the smooth tracking generation. Action and perception are tightly coupled in a bidirectional way: perception triggers an action and the output of action changes the perception. Meanwhile, the robot motion information is fed forward, inspired by the VOR mechanism, to stabilize smooth tracking.

The paper is organized as follows. Section 2 introduces bionic issues and design of our bionic SPM. Section 3 proposes visual tracking based on symmetric Kullback-Leibler metric *spatiograms*. Our bionic eye plant control system is described in Section 4. Experimental results are shown in Section 5. Section 6 presents our conclusion.

2. Design of Human-Eye-Inspired PTZ Platform

2.1. Human-Eye’s Movement Mechanism. Each eye is controlled by three complementary pairs of extraocular muscles, as shown in Figure 1(a). The movement of each eye involves rotating the globe of the eye in the socket of the skull. Because of minimal translation during its movement, the eye can be regarded as a spherical joint with an orientation defined by three axes of rotation (horizontal, vertical, and torsional). But in our implementation and development of a simulator, we view eye movement with no translation for simplicity.

The medial rectus turns eye inward and, thus, lateral rectus outward. Therefore, they form a pair to control the horizontal position of the eye. In contraction to the pair of medial rectus and lateral rectus, the actions of the other two

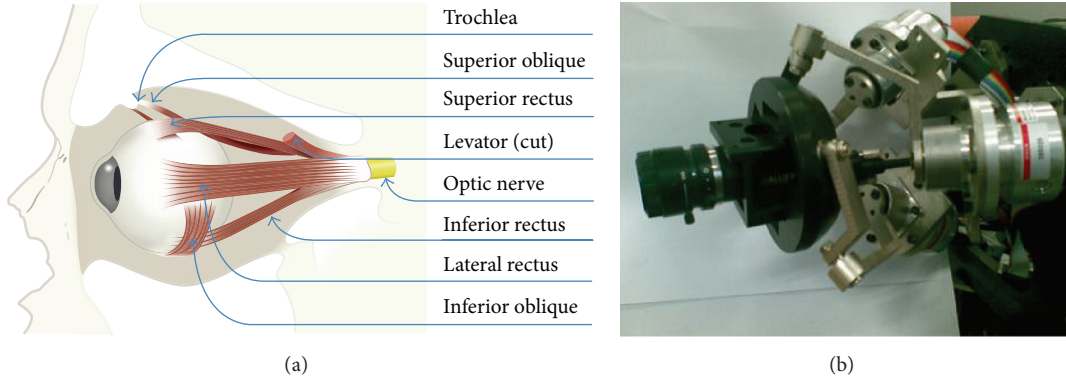


FIGURE 1: Development of bionic eye plant. (a) Muscles of the eye. Six muscles, arranged in three pairs, control the movements of the eye as shown here in a cutaway view of the eye in its socket or orbit. (b) Structure of our SPM prototype.

pairs of muscles are more complex. When the eye is centered in the orbit, the primary effect of the superior and inferior rectus is to rotate up or rotate down the eye. However, when the eye is deviated horizontally in the orbit, these muscles also contribute to torsion, the rotation of the eye around the line of sight that determines the orientation of images on the retina.

The primary effect of the superior and inferior obliques is to turn eyes downward and upward when the eye does not deviate from horizontal position. So do superior rectus and inferior rectus. In addition, these muscles also determine the vertical orientation of the eye.

Smooth pursuit eye movements slowly rotate the eyes to compensate for any motion of the visual target and thus act to minimize the blurring of the target's retinal image that would otherwise occur. We implement smooth target tracking, continuously adjusted by visual feedback about the target's image (retinal image).

Kinematic characteristics of SPM and the mechanics of eye movements are very similar [29]. Both have a 3-DOF spherical movement and rotating globe is the center of the sphere. SPM also has a compact structure, excellent dynamic performance, and high accuracy, so 3-DOF SPM is in line with the structural design of the bionic eye to replicate the eye movement.

The eyeball is seen as a sphere, with a rotation center when it rotates. Inspired by the mechanics of eye movements and active robotic vision, we presented a new bionic eye prototype based on SPM, which is made up of an eye-in-hand system as shown in Figure 1(b).

Because the eye is free to rotate in three dimensions, eyeballs can keep retinal images stable in the fovea when they track moving target. In our work, we proposed two main points about structural requirements inspired by the human eyes: (1) camera center must be located at the center of "eyeball" to ensure that the angle of image planes between two different positions keeps identical with the rotation of "eyeball"; (2) in the process of eye movement, any mechanical component except the "eyeball" cannot exceed the plane of the center of the sphere as much as possible to ensure that when the movement of the "eyeball" occurs, they do not block the sight of the camera and do not interfere with the robot face.

2.2. Oculomotor Plant Compensation of VOR. In the human-eye VOR, a signal from the vestibular system related to head velocity, which is encoded by semicircular ducts, is used to drive the eyes in the opposite direction to the head movement. The VOR operates in feedforward mode and as such requires calibration to ensure accurate nulling of head movement. The simplicity of this "three-neuron arc," together with the relatively straightforward mechanics of the eye plant, has long made the VOR an attractive model for experimental and computational neuroscientists seeking to understand cerebellar function. To abolish image motion across the retina, the vestibular signal must be processed by neural circuitry which compensates for the mechanical properties of the oculomotor plant. The VOR is therefore a particular example of motor plant compensation. Horizontal and vertical and angular and linear head movement motivate the appropriate combinations of six extraocular muscles in three dimensions.

2.3. Kinematics of Human-Eye-Inspired PTZ Platform. Spherical parallel mechanism consists of an upper layer and a base, connected by three pairs of identical kinematic subchains as shown in Figure 2. In each chain, there is one fixed revolute joint z_i and two free revolute joints x_i and y_i connecting the proximal link to the distal link and the distal link to the upper layer, respectively. The axes of all revolute joints intersect at a common point O which is referred to as the rotational center. The plane passing through the point O and becoming parallel with the base is called the sphere center plane, also seen as Listing's plane of eyeball. $\alpha_1, \alpha_2, \beta_1, \beta_2$, and η are the parameters of this mechanism, where α_1 and α_2 are the structural angle of the lower link and upper link, β_1 and β_2 are the half-cone angle of the upper platform and the base, and η is the structural torsion angle of initial state of the upper platform and the base, namely, the initial torsion angle.

Figure 2 demonstrates the kinematics of our SPM platform, and the kinematic equation of the SPM is given by [30]

$$\dot{\theta} = J(\alpha_1, \alpha_2, \beta_1, \beta_2, \eta, \phi_x, \phi_y, \phi_z) \omega, \quad (1)$$

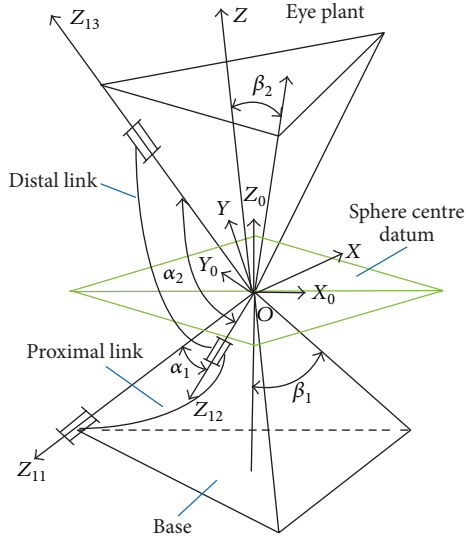


FIGURE 2: Kinematic sketch of a spherical parallel manipulator.

where $\dot{\theta} = (\dot{\theta}_1, \dot{\theta}_2, \dot{\theta}_3)$ is the angular velocity vector input of the motor, $\omega = (\omega_x, \omega_y, \omega_z)$ is the angular velocity vector output of the upper platform, and J is the Jacobian matrix which is decided by the mechanical parameter ($\alpha_1, \alpha_2, \beta_1, \beta_2$, and η) and the eyeball posture (ϕ_x, ϕ_y , and ϕ_z). α_1 and α_2 are the structural angles of lower link and upper link, respectively. β_1 and β_2 are the angles of base and upper platform. The proposed PTZ platform has similar kinematics to the human eye, as shown in Figure 3.

3. SKL-Based Particle Filter Visual Tracking

3.1. Spatial Histograms: Spatiogram. Color histogram is one of the common target models which is just a statistic of different colors in the entire picture in proportion without concern for spatial location of each color. Therefore, it is not rather sensitive to rotation but suitable for nonrigid or prone to deformation modeling target objects. Targets based on this model are vulnerable to backgrounds which have similar color distribution or other interference, thereby causing the target tracking failure. In this paper, we improve the particle filter algorithm based on a new target model, spatiogram [31], which adds the pixel coordinate information to the traditional color histogram and uses SKL metric. The second-order spatiogram can be described as follows:

$$h(b) = \{n_b, \mu_b, \Sigma_b\}, \quad b = 1, \dots, B, \quad (2)$$

where B is the total number of the intervals and $\{n_b, \mu_b, \Sigma_b\}$ is the probability of each interval, coordinate mean, and covariance matrix, respectively. They can be calculated using the formula as follows:

$$n_b = \frac{1}{N} \sum_{j=1}^N \delta_{jb},$$

$$\mu_b = \frac{1}{N * n_b} \sum_{j=1}^N \mathbf{x}_j \delta_{jb},$$

$$\Sigma_b = \frac{1}{N * n_b - 1} \sum_{j=1}^N (\mathbf{x}_j - \mu_b) (\mathbf{x}_j - \mu_b)^T \delta_{jb}. \quad (3)$$

B is the total number of pixels within the target area, $\mathbf{x}_j = [x_j, y_j]^T$ is the coordinate position of the j th pixel, and $\delta_{jb} = 1$ denotes that the j th pixel is quantized to the b th interval, while $\delta_{jb} = 0$ indicates that the j th pixel is quantized to other intervals.

3.2. SKL-Based Particle Filter. In order to apply the spatiogram to target tracking, we need to select a method to measure the similarity metrics of the spatial histogram between the targets and the candidate targets. We select the SKL-based coefficient of similarity metrics to measure the similarity of the target spatiogram $h(b) = \{n_b, \mu_b, \Sigma_b\}$ and candidate target spatiogram $h'(b) = \{n'_b, \mu'_b, \Sigma'_b\}$.

Given a spatiogram $h(b) = \{n_b, \mu_b, \Sigma_b\}$, we use a Gaussian distribution to describe the spatial distribution of all the pixels in each section. The distribution of the b th section can be described as

$$f_b(\mathbf{x}) = \frac{1}{2\pi |\Sigma_b|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x} - \mu_b)^T \Sigma_b^{-1} (\mathbf{x} - \mu_b) \right], \quad (4)$$

where μ_b is the mean value of all coordinates of the pixels of the b th interval and Σ_b is the mean covariance matrix of all coordinates of the pixels of the b th interval.

The KL distance between the Gaussian distribution $f_b(\mathbf{x})$ and the Gaussian distribution $f'_b(\mathbf{x})$ can be obtained by a closed form solution which is calculated using the following formula:

$$\text{KL}(f_b \parallel f'_b) = \frac{1}{2} \left[\log \left(\frac{|\Sigma'_b|}{|\Sigma_b|} \right) + \text{Tr} \left((\Sigma'_b)^{-1} \Sigma_b \right) - d \right. \\ \left. + (\mu_b - \mu'_b)^T (\Sigma'_b)^{-1} (\mu_b - \mu'_b) \right], \quad (5)$$

where d is the spatial dimension (for spatiogram, $d = 2$). Similarly, we can get the KL distance between the Gaussian distribution $f'_b(\mathbf{x})$ and the Gaussian distribution $f_b(\mathbf{x})$:

$$\text{KL}(f'_b \parallel f_b) = \frac{1}{2} \left[\log \left(\frac{|\Sigma_b|}{|\Sigma'_b|} \right) + \text{Tr} \left(\Sigma_b^{-1} \Sigma'_b \right) - d \right. \\ \left. + (\mu'_b - \mu_b)^T \Sigma_b^{-1} (\mu'_b - \mu_b) \right]. \quad (6)$$

The SKL distance of the two Gaussian distributions of $f_b(\mathbf{x})$ and $f'_b(\mathbf{x})$ is

$$\text{SKL}(f_b \parallel f'_b) = \frac{1}{2} \left[\text{KL}(f_b \parallel f'_b) + \text{KL}(f'_b \parallel f_b) \right],$$

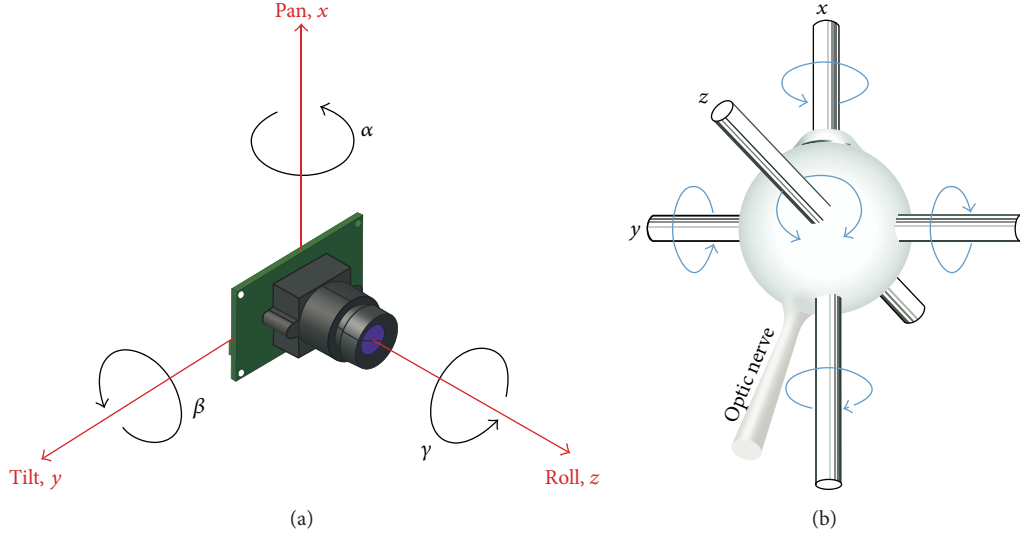


FIGURE 3: Analogue of camera rotation and eye movement. Our SPM prototype has similar kinematics to human eye.

$$\begin{aligned}
 & \text{SKL}(f_b \parallel f'_b) \\
 &= \frac{1}{4} \left[\text{Tr} \left((\Sigma'_b)^{-1} \Sigma_b \right) + \text{Tr} \left(\Sigma_b^{-1} \Sigma'_b \right) - 2d \right] \\
 &+ \frac{1}{4} \left[(\boldsymbol{\mu}_b - \boldsymbol{\mu}'_b)^T \left(\Sigma_b^{-1} + (\Sigma'_b)^{-1} \right) (\boldsymbol{\mu}_b - \boldsymbol{\mu}'_b) \right].
 \end{aligned} \tag{7}$$

Generally, the ranges of the similarity are $[0, 1]$, and the similarity ψ_b of each pair of intervals on the spatiogram can be described as

$$\psi_b = \exp \left[-\text{SKL}(f_b, f'_b) \right]. \tag{8}$$

Thus, the similarity of the spatiogram based on SKL distance can be calculated as

$$\begin{aligned}
 \rho(h, h') &= \sum_{b=1}^B \sqrt{n_b n'_b} \psi_b \\
 &= \sum_{b=1}^B \sqrt{n_b n'_b} \exp \left[-\text{SKL}(f_b, f'_b) \right].
 \end{aligned} \tag{9}$$

According to (11), we can get

$$\begin{aligned}
 \rho(h, h') &= \sum_{b=1}^B \sqrt{n_b n'_b} \\
 &\cdot \exp \left[-\frac{1}{4} \left(\text{Tr} \left(\Sigma_b^{-1} \Sigma'_b \right) + \text{Tr} \left(\Sigma'_b^{-1} \Sigma_b \right) - 2d + 0 \right) \right] \\
 &= \sum_{b=1}^B n_b \exp \left[-\frac{1}{4} (d + d - 2d) \right] = \sum_{b=1}^B n_b = 1.
 \end{aligned} \tag{10}$$

This indicates that the similarity measure of symmetric spatiogram based KL distance ensures that the object has the most similarity to the target.

4. Image-Based Feedback and Dynamic Compensation Eye Plant Control

4.1. Visual Feedback Scheme. When the target is identified in the image, the visual feedback tracking control strategy is proposed to control the bionic eye plant mechanism to minimize a tracking error function, which is also called eye-in-hand visual servoing [32, 33]. Since the relative distance between the SPM and the moving target is large, if the error function is defined in any 3D reference coordinate frame, coarse estimation of the relative pose between the SPM and the moving target may cause the moving target to fall out of the visual field, while adjusting the SPM servo mechanism, and also affect the accuracy of the pose reached after convergence. In our project, to make tracking control more robust and stable, we define a tracking error function in the visual sensor frame, which is given by [32, 33]

$$\mathbf{e}(t) = \mathbf{s}(t) - \mathbf{s}^*, \tag{11}$$

where $\mathbf{s}(t)$ and \mathbf{s}^* are the measured and desired locations of the centroid of the tracked moving target with respect to the image plane, respectively. In our work, we set $\mathbf{s}^* = [0, 0]^T$, a constant, which is the centroid of the captured image.

4.2. Camera Calibration of Human-Eye-Inspired PTZ Platform. Based on the PTZ visual system, the coordinate system is established as shown in Figure 3. Assume the motion of the object is unknown; how do we control motion of the PTZ platform so that the projection of the moving object is fixed at the center of the image plane, with full consideration of the dynamic effects of the “eyeball”? To make z-axis of the camera coordinate system coincide with the target by adjusting the posture of the camera, we have to compensate the offset angle between the camera and the target. We employ a pinhole camera model to obtain a more accurate camera projection.

Following the general pinhole camera model, the intrinsic parameter model equation of the camera is given by

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{x_c}{z_c} \\ \frac{y_c}{z_c} \\ 1 \end{bmatrix}, \quad (12)$$

where (u, v) denotes the image coordinate of the target in the image coordinate system. (u_0, v_0) are the coordinates of

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \gamma & -\sin \gamma \\ 0 & \sin \gamma & \cos \gamma \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ z_T \end{bmatrix}. \quad (13)$$

Finally, we can deduce the angle offset between the target and camera's line of sight:

$$\alpha = \arctan \left(\frac{u}{f_x} - \frac{u_0}{f_x} \right), \quad (14)$$

$$\beta = \arctan \left(\frac{f_x}{f_y} \sqrt{\frac{(v - v_0)^2}{f_x^2 + (u - u_0)^2}} \right).$$

In our implementation, the camera center is located at the center of "eyeball" so that the angle of image planes between two different positions keeps identical with the rotation of "eyeball." The 3-DOF SPM satisfies the principles of eyeball movement; a camera can be mounted in the SPM and actively oriented (horizontal, vertical, and torsional) around its x -axis, y -axis, and z -axis, respectively. We, considering minimal translation during eye's movement, implement our eye plant system with no translation for simplicity. So our visual tracking strategy is applicable to all SPMs with no translation.

In the visual tracking section, we give how to determine the position of the moving object. The relative position determines our visual tracking strategy. Eye rotation about the vertical "z-axis" is controlled by the lateral and medial rectus muscles, which results in eye movements to left or right. Rotation about the transverse "y-axis" is controlled by the superior and inferior rectus muscles, which elevates and depresses the eye. Finally, rotations about the anteroposterior "x-axis" result in counterclockwise as well as upward and downward eye motion. See Figures 1(a) and 1(b). Our model receives visually guided signal to control eye plant; see (14). Meanwhile, the robot motion information is fed forward into control loop. Our whole bioinspired tracking system is illustrated in Figure 4. It is known that the VOR is basically driven by the signals from vestibular apparatus in the inner ear. The semicircular canals (SCs) detect the head rotation and drive the rotational VOR; on the other hand, the otoliths detect the head translation and drive the translational VOR. Anatomists and physiologists tend to engage in the VOR

the principal point. (x_c, y_c, z_c) is the target coordinate in the camera coordinate system. f_x is the scale factor in the x -coordinate direction, and f_y is the scale factor in the y -coordinate direction.

In order to keep the target tracked in the center of the field, we need to make the target lie on the optical axis. The location of the target which passes through the optical axis is represented by $(0, 0, z_T)$, where z_T is the distance between the camera and the target. The orientation is

as a simple neural system mediated by a three-neuron arc and displaying a distinct function. Starting in the vestibular system, SCs get activated by head rotation and send their impulses via the vestibular nerve through brainstem neurons and end in oculomotor plant. Here, we use IMU to acquire pose change of eye from the robot.

When the robot works in the bumpy environment, rigid bumps and pulse jitter cause the occurrence of significant turbulence with high frequency and posture change with lower frequency. Therefore, the motion information of the robot is acquired and fed forward into the controller to compensate the external disturbance. In [34], an active compensation model of visual error is proposed according to the principle of VOR in (15). Here, we use our proposed bioinspired controller to compensate motion disturbance caused by bumpy jitter. Hence,

$$E(s) = H(s) \frac{-\alpha T_v T_n s^2}{(sT_v + 1)((sT_n + 1))} + e(s) \left[(\lambda s + \gamma) + \frac{ke^{-ls}}{sT_q} \right] \frac{T_n}{sT_n + 1}, \quad (15)$$

where $e(s) = -(H(s) + E(s))$ is slide error of retina, $H(s)$ denotes the rotation angle of head, and $E(s)$ means the rotation angle of eyeball. α , λ , and γ represent the gains of the velocity signal of head rotation, the velocity signal of retina slide, and the spike of nerve fibers caused by the displacement of retina, respectively. k is the compensation weight value of flocculus caused by error signal of retina. In our system, α , λ , and γ are equal to 1 and $k = 2.5$. Combining position compensation with speed compensation of eyeball, our system is used to build a smooth tracking system.

5. Experiments and Results

To prove that the proposed image-based feedback tracking system based on our developed eye-in-hand prototype is able to orient a camera with the required orientation changes,

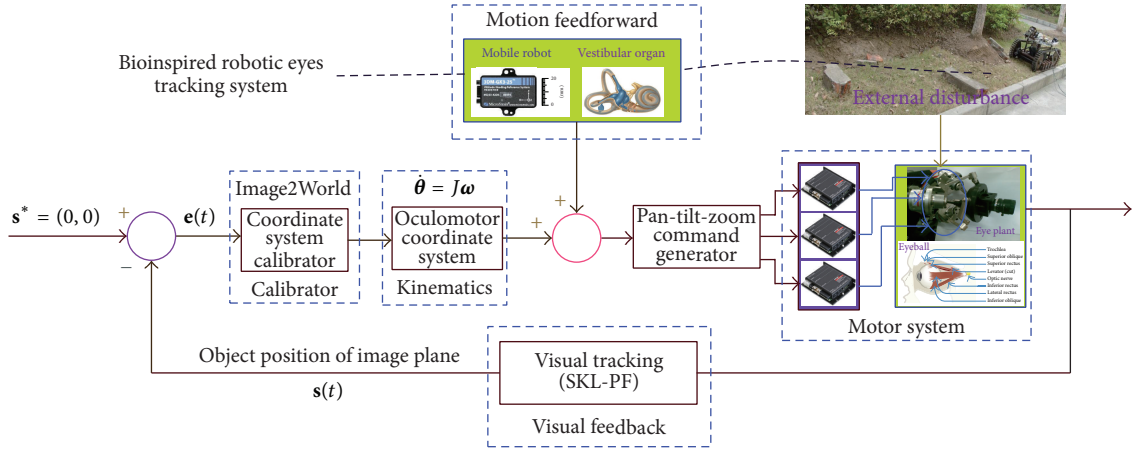


FIGURE 4: Tracking control scheme. Camera means human's eyeball; IMU means canals.

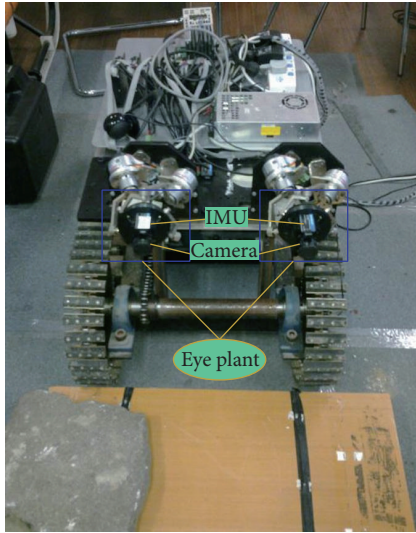


FIGURE 5: Experimental platform based on Tracked Mobile Robot.

especially its dynamic disturbance resistance capability and SPM-based structural dexterity, closed-loop control experiments were performed. We design an experimental platform based on a tracked robot, as shown in Figure 5. A variety of obstacles are placed on the tracked robot's path to simulate a real harsh environment. We introduced the used joint space control architecture in [35]. In the chosen control approach, the desired camera orientation is transformed to linear actuator set points using the inverse kinematics. Thus, here only a brief overview of the architecture and exemplary control results are presented.

To measure angular velocities of "eyeball" in three axes, we employ the attitude sensor 3DM-GX-25TM. The device offers a range of output data quantities from fully calibrated inertial measurements (acceleration, angular rate, and deltaAngle and deltaVelocity vectors) to computed orientation estimates, including pitch, roll, and heading (yaw) or

rotation matrix. All quantities are fully temperature compensated and are mathematically aligned to an orthogonal coordinate system.

In addition, the image information is gained by using a high-speed camera (Guppy F-033C), which is connected to the IEEE 1394 card installed in a PC with Intel Core CPU which acquires the video signal. The camera is an ultracompact, inexpensive VGA machine vision camera with CCD sensor (Sony ICX424). At full resolution, it runs up to 58 fps. We employ a pinhole camera model to obtain a more accurate camera projection. Camera calibration was repeated ten times to seek an approximate camera calibration matrix $K = \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ u_0 & v_0 & 1 \end{bmatrix}$. A camera is calibrated using chessboard [36]. Here, we employ a pinhole camera model to obtain a more accurate camera projection. Following the general pinhole camera model, the parameters contained in K are called the internal camera parameters or the internal orientation of the camera. See [37] for more details.

Figure 4 shows our tracking control scheme. We implemented smooth target tracking (smooth pursuit) to ensure the target located in the field of view, continuously adjusted by visual feedback about the target's image. Image is captured from camera (retinal image) and IMU measures the robot body's movement to compensate dynamic disturbance.

Supposing that we do not know the motion of the tracked object, how do we control the motion of the "eyeball" to ensure that the moving object is fixed at the centroid of the image plane?

In the process of tracking moving target, the tracking algorithm should be robust to appearance variations introduced by occlusion, illumination changes, and pose variations. In our library environment, the proposed algorithm can relocate the target when object appearance changes due to illumination, scale, and pose variations. Once the moving target is located, the "eyeball" should keep images stable in the field of view (center of image). That is, target position fluctuates at zero. See Figures 6 and 7. Figure 6 gives some snapshots of tracking results and demonstrates that

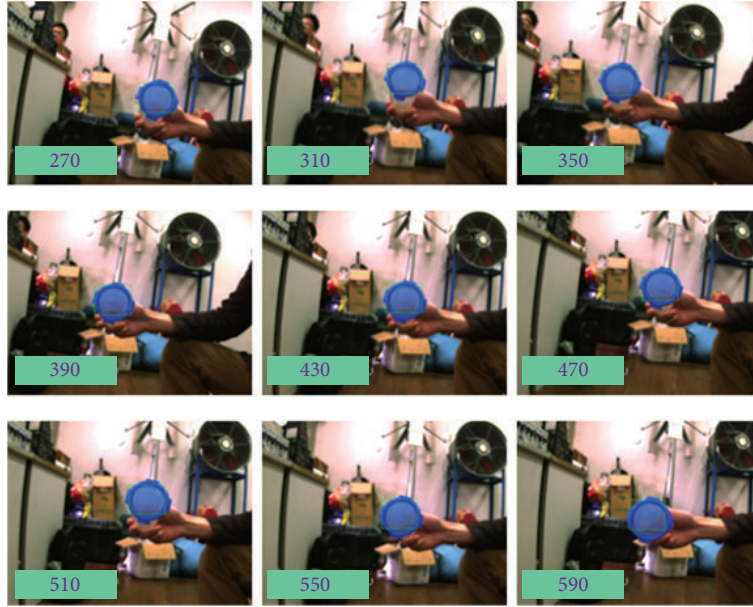


FIGURE 6: Moving object tracked in the laboratory.

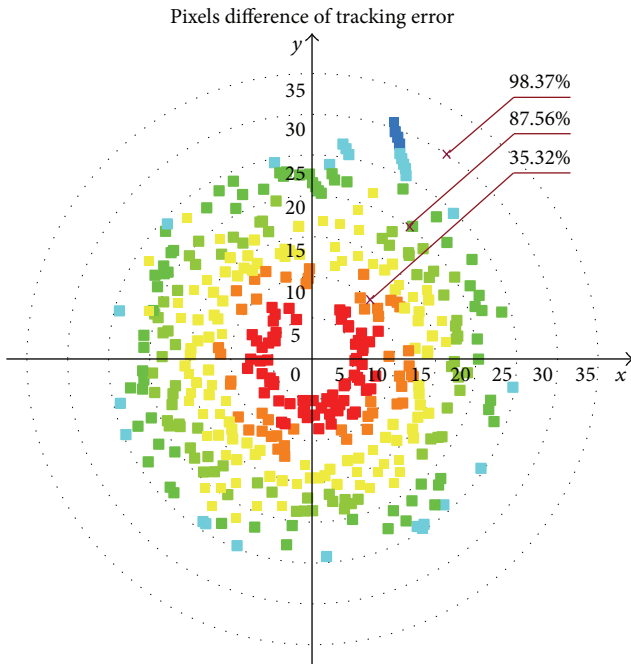


FIGURE 7: Pixel difference of x and y directions.

the moving target is located in the field of view. Meanwhile, extensive experiments are conducted to perform bumpy-resist capability. Figure 7 illustrates the pixel difference in x and y direction. Smaller eyeball errors accompanying larger postural changes can be good proofs of good bumpy-resist capability and VOR function.

Figures 7 and 8 show the performance of the proposed tracking system on the tracked robot running across bumpy environment. The result of statistics shows that 98.37% of

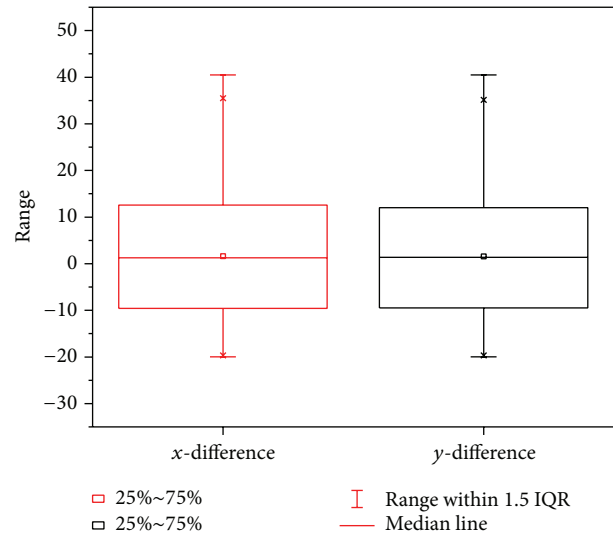


FIGURE 8: Statistics of x and y direction difference.

tracking errors, including x and y direction difference, have fallen into the range of <30 pixels, as shown in Figure 7. The statistics of x and y direction pixel difference are demonstrated in Figure 8. In our test, as the tracked robot platform travels through the rough ground full of obstacles, rigid bumps and pulse jitter cause the occurrence of significant turbulence with high frequency and the oscillatory posture changes with lower frequency, which makes tracking effect slightly larger than the data recorded in the literature, such as [13, 14]. But our experiments are established under relatively harsh environmental conditions, and the effect achieved is objective. Tracking effects still stay in a controllable range like the above situation. Apparently, this indicates that the system has robustness.

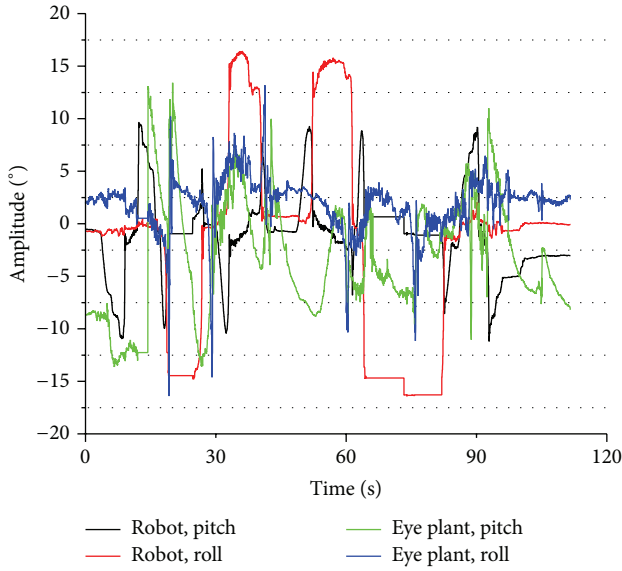


FIGURE 9: Experimental results of robot's and eye plant's pose.

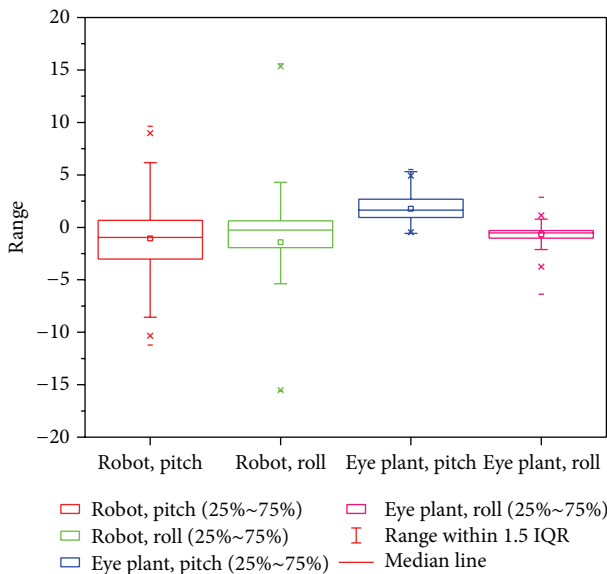


FIGURE 10: Statistics of robot's and eye plant's pose.

In our actual situation, we install three IMUs on the tracked robot and eye plants to measure the pose changes. We recorded the angle variances to validate the system bumpy-resist capability that the eyeball moves on the opposite direction according to position compensation and velocity compensation when the tracked robot's pose changes. In other words, the robot pose variance information is fed forward into controller to form a head-eye coordination system. Figures 9 and 10 show the experimental results of tracked robot's and eye plant's pose changes on the tracked robot in bumpy environment. In addition, the large tracking errors happen when the robot encounters instantaneous postural changes. Nonetheless, quick returns to lower errors of eyeball verify good robustness of the bionic visual tracking system

and high dexterity of the SPM-based bionic eye. Obviously, these variances reflect good stability of the tracking system.

6. Conclusion

To accurately replicate the human vision system, we presented a 3-DOF "eyeball" in the directions of horizontal, vertical, and torsional axes according to the mechanics of eye movements. Thus, an image-based visual feedback tracking system is presented to minimize a tracking error function, capable of tracking moving target. More specifically, the proposed real-time moving target tracking algorithm utilizes spatial histograms and symmetric Kullback-Leibler metric integrated into particle filtering framework to achieve automatic moving target identification and gaze stabilization. Meanwhile, the robot motion information is fed forward to develop an adaptive smooth tracking controller bio-inspired by the VOR mechanism. The experimental results demonstrate that our algorithm is effective and robust in dealing with moving object tracking and can always keep the target at the center of the camera to avoid tracking failure. Furthermore, as the tracked robot platform travels through the rough ground full of obstacles, rigid bumps and pulse jitter cause the occurrence of significant turbulence with high frequency and the oscillatory posture changes with lower frequency. Tracking effects still stay in a controllable range and this indicates that the system has bumpy-resist capability.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was partly supported by the National Natural Science Foundation of China (nos. 61233010 and 61305106), the Nature Science Foundation of Shanghai (no. 13ZR1454200), and the Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (no. ICT1531).

References

- [1] B. Jung and G. S. Sukhatme, "Real-time motion tracking from a mobile robot," *International Journal of Social Robotics*, vol. 2, no. 1, pp. 63–78, 2010.
- [2] M. Hwangbo, J.-S. Kim, and T. Kanade, "Gyro-aided feature tracking for a moving camera: fusion, auto-calibration and GPU implementation," *The International Journal of Robotics Research*, vol. 30, no. 14, pp. 1755–1774, 2011.
- [3] M. Hwangbo, J.-S. Kim, and T. Kanade, "Inertial-aided KLT feature tracking for a moving camera," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '09)*, pp. 1909–1916, St. Louis, MO, USA, October 2009.
- [4] J. Park, W. Hwang, H. Kwon, K. Kim, and D.-I. D. Cho, "A novel line of sight control system for a robot vision tracking system, using vision feedback and motion-disturbance feedforward compensation," *Robotica*, vol. 31, no. 1, pp. 99–112, 2013.

- [5] Z. Jia, A. Balasuriya, and S. Challa, "Sensor fusion-based visual target tracking for autonomous vehicles with the out-of-sequence measurements solution," *Robotics and Autonomous Systems*, vol. 56, no. 2, pp. 157–176, 2008.
- [6] J. D. Hol, T. B. Schön, H. Luinge, P. J. Slycke, and F. Gustafsson, "Robust real-time tracking by fusing measurements from inertial and vision sensors," *Journal of Real-Time Image Processing*, vol. 2, no. 2-3, pp. 149–160, 2007.
- [7] A. Lenz, T. Balakrishnan, A. G. Pipe, and C. Melhuish, "An adaptive gaze stabilization controller inspired by the vestibulo-ocular reflex," *Bioinspiration and Biomimetics*, vol. 3, no. 3, Article ID 035001, 2008.
- [8] T. Shibata and S. Schaal, "Biomimetic gaze stabilization based on feedback-error-learning with nonparametric regression networks," *Neural Networks*, vol. 14, no. 2, pp. 201–216, 2001.
- [9] H. Xu, Y. Xu, H. Fu, Y. Xu, X. Z. Gao, and K. Alipour, "Coordinated movement of biomimetic dual PTZ visual system and wheeled mobile robot," *Industrial Robot*, vol. 41, no. 6, pp. 557–566, 2014.
- [10] O. Avni, F. Borrelli, G. Katzir, E. Rivlin, and H. Rotstein, "Scanning and tracking with independent cameras—a biologically motivated approach based on model predictive control," *Autonomous Robots*, vol. 24, no. 3, pp. 285–302, 2008.
- [11] J. Law, P. Shaw, and M. Lee, "A biologically constrained architecture for developmental learning of eye-head gaze control on a humanoid robot," *Autonomous Robots*, vol. 35, no. 1, pp. 77–92, 2013.
- [12] S. Xie, J. Luo, Z. Gong, W. Ding, H. Zou, and X. Fu, "Biomimetic control of pan-tilt-zoom camera for visual tracking based-on an autonomous helicopter," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '2007)*, pp. 2138–2143, San Diego, Calif, USA, November 2007.
- [13] L. Vannucci, N. Cauli, E. Falotico, A. Bernardino, and C. Laschi, "Adaptive visual pursuit involving eye-head coordination and prediction of the target motion," in *Proceedings of the 14th IEEE-RAS International Conference on Humanoid Robots (Humanoids '14)*, pp. 541–546, IEEE, Madrid, Spain, November 2014.
- [14] E. Falotico, D. Zambrano, G. G. Muscolo, L. Marazzato, P. Dario, and C. Laschi, "Implementation of a bio-inspired visual tracking model on the iCub robot," in *Proceedings of the 19th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN '10)*, pp. 564–569, IEEE, Viareggio, Italy, September 2010.
- [15] L. Jamone, M. Fumagalli, G. Metta, L. Natale, F. Nori, and G. Sandini, "Machine-learning based control of a human-like tendon-driven neck," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '10)*, pp. 859–865, Anchorage, Alaska, USA, May 2010.
- [16] F. Nori, L. Jamone, G. Sandini, and G. Metta, "Accurate control of a human-like tendon-driven neck," in *Proceedings of the 7th IEEE-RAS International Conference on Humanoid Robots*, pp. 371–378, IEEE, Pittsburgh, Pa, USA, November-December 2007.
- [17] N. G. Tsagarakis, G. Metta, G. Sandini et al., "iCub: the design and realization of an open humanoid platform for cognitive and neuroscience research," *Advanced Robotics*, vol. 21, no. 10, pp. 1151–1175, 2007.
- [18] J. Leitner, S. Harding, M. Frank, A. Förster, and J. Schmidhuber, "An integrated, modular framework for computer vision and cognitive robotics research (icVision)," in *Biologically Inspired Cognitive Architectures*, vol. 196 of *Advances in Intelligent Systems and Computing*, pp. 205–210, Springer, 2013.
- [19] X.-Y. Wang, Y. Zhang, X.-J. Fu, and G.-S. Xiang, "Design and kinematic analysis of a novel humanoid robot eye using pneumatic artificial muscles," *Journal of Bionic Engineering*, vol. 5, no. 3, pp. 264–270, 2008.
- [20] Y.-C. Lee, C.-C. Lan, C.-Y. Chu, C.-M. Lai, and Y.-J. Chen, "A pan-tilt orienting mechanism with parallel axes of flexural actuation," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 3, pp. 1100–1112, 2013.
- [21] C.-C. Lan, Y.-C. Lee, J.-F. Jiang, Y.-J. Chen, and H.-Y. Wei, "Design of a compact camera-orienting mechanism with flexural pan and tilt axes," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '11)*, pp. 1962–1967, San Francisco, Calif, USA, September 2011.
- [22] G. Cannata and M. Maggiali, "Models for the design of bioinspired robot eyes," *IEEE Transactions on Robotics*, vol. 24, no. 1, pp. 27–44, 2008.
- [23] J. Gu, M. Meng, A. Cook, and M. G. Faulkner, "A study on natural movement of artificial eye implant," *Robotics and Autonomous Systems*, vol. 32, no. 2, pp. 153–161, 2000.
- [24] F. Colonnier, A. Manecy, R. Juston et al., "A small-scale hyperacute compound eye featuring active eye tremor: application to visual stabilization, target tracking, and short-range odometry," *Bioinspiration & Biomimetics*, vol. 10, no. 2, Article ID 026002, 2015.
- [25] T. Villgratner, E. Schneider, P. Andersch, and H. Ulbrich, "Compact high dynamic 3 DoF camera orientation system: development and control," *Journal of System Design and Dynamics*, vol. 5, no. 5, pp. 819–828, 2011.
- [26] T. Villgratner and H. Ulbrich, "Optimization and dynamic simulation of a parallel three degree-of-freedom camera orientation system," in *Proceedings of the 23rd IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems (IROS '10)*, pp. 2829–2836, Taipei, Taiwan, October 2010.
- [27] C. M. Gosselin, E. St. Pierre, and M. Gagné, "On the development of the agile eye," *IEEE Robotics & Automation Magazine*, vol. 3, no. 4, pp. 29–37, 1996.
- [28] Y.-B. Bang, J. K. Paik, B.-H. Shin, and C. Lee, "A three-degree-of-freedom anthropomorphic oculomotor simulator," *International Journal of Control, Automation and Systems*, vol. 4, no. 2, pp. 227–235, 2006.
- [29] S. Refaat, J. M. Hervé, S. Nahavandi, and H. Trinh, "Two-mode overconstrained three-DOFs rotational-translational linear-motor-based parallel-kinematics mechanism for machine tool applications," *Robotica*, vol. 25, no. 4, pp. 461–466, 2007.
- [30] C. Li, S. Xie, H. Li, D. Wang, and J. Luo, "Design of bionic eye based on spherical parallel mechanism with optimized parameters," *Robot*, vol. 6, article 014, 2010.
- [31] S. T. Birchfield and S. Rangarajan, "Spatiograms versus histograms for region-based tracking," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 2, pp. 1158–1163, IEEE, June 2005.
- [32] F. Chaumette and S. Hutchinson, "Visual servo control. I. Basic approaches," *IEEE Robotics and Automation Magazine*, vol. 13, no. 4, pp. 82–90, 2006.
- [33] F. Chaumette and S. Hutchinson, "Visual servo control. II. Advanced approaches [Tutorial]," *IEEE Robotics and Automation Magazine*, vol. 14, no. 1, pp. 109–118, 2007.
- [34] H. Li, J. Luo, C. Li, L. Li, and S. Xie, "Active compensation method of robot visual error based on vestibulo-ocular reflex," *Jiqiren/Robot*, vol. 33, no. 2, pp. 191–197, 2011.

- [35] C. Li, S. Xie, H. Li, J. Miao, Y. Xu, and J. Luo, "System design and study on bionic eye of spherical parallel mechanism based on attitude closed-loop control," *Jiqiren/Robot*, vol. 33, pp. 354–359, 2011.
- [36] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [37] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, UK, 2003.