Optimal Trajectory Planning of Eye-head Gaze Shifts for Bionic Robot Vision System on Orchard Environment Perception

Jianping Ma, Yuanhao Cheng and Sun’an Wang

School of Mechanical Engineering, Xi’an Jiaotong University, No.28, Xianning West Road, Xi’an, Shaanxi province, China
Email: 1821582672@qq.com; cyhv2006@126.com; sawang@xjtu.edu.cn;

Abstract. To make orchard robots obtain environmental information quickly and smoothly, this paper studies the neuro-mechanism of primates’ gaze shifts. Meanwhile, according to the optimal control theory, this paper considers time and energy as factors of formulating cost function and then analyses kinematic relationship between eye and head during gaze shifts to find constraint conditions. Solving the two-point boundary value differential equations of optimal control, this paper can extract the best control signal of coordinated eye and head movements. By building a multi-degree-of-freedom robot vision system corresponding to the biological model, this paper simulates primates’ gaze shifts and verifies the feasibility and validity of the method proposed.

1. Introduction
Vision is an important form for orchard robots to obtain information such as roads, trees, fruits and so on. However, structure design and control strategy selection of robot vision system greatly affect its stability and efficiency in an orchard environment. Building intelligent bionic eyes by using the neuro-mechanism of biological model provides a good solution for orchard robot vision. Research [1] found that primates with the highest degree of evolution on vision system all have two eyes. Additionally, when observing targets, they have to move both eyes and head to transfer the gaze point. In this paper, we simulate primates’ eye and head movements during gaze shifts to make the orchard robot vision system move like primates.

To simulate the process of primates’ gaze shifts, scholars have done a lot of research from two aspects: building mathematical models and optimal control. Research [2] analyses the change of input and output signals of neurons and perceptron involved in the process of primates’ gaze shifts and builds a corresponding mathematical model. Research [3] proposes a segmented eye and head coordination control strategy based on the horizontal direction of gaze point. They all study the gaze shifts by building simplified mathematical model. Research [4] studies the relationship between eye and head movements during gaze shifts, and proposes an optimal control strategy by minimizing the cost force or energy. Research [5] proposes an optimal adaptive control model of saccadic process. Research [6] studies natural eye and head movements by minimizing the impact of noise. Research [7] proposes an optimal control strategy by minimizing the integral of absolute error in control process. They all studies the gaze shifts by optimal control. However, the selected factors for building cost function are single. Meanwhile, they are also absent from practical engineering application.

Compared with building mathematical models, optimal control strategy doesn’t consider the details of neurons or perceptron. It establishes the cost function as a selection mechanism to extract the best
input-output relationship from a large number of experimental data. Therefore, this method can improve the efficiency of bionic vision system while retaining biological characteristics. But at present, eye and head coordination problems mostly stay at the level of computer simulation. What’s more, the cost function is selected by single factor. Their work focuses more on academic research than on practical engineering application. So, based on biological mathematical model [2, 9-11] and practical engineering requirements, this paper selects multiple related factors to formulate cost function and solves the optimal control differential equations.

2. Biological Model

The factors that influence cost function’s building include time, accuracy, force, energy, stability, noise, impulse and jerk [12]. For actual engineering system, real-time control is the most important. And reducing energy consumption can help orchard robots carry out a longer-lasting task. So, in this paper, we consider time and energy as the main factors building cost function. Meanwhile, stability, jerk and impulse are taken as the constraint conditions. The representation of the time part: 
\[ J = \int_0^T dt = t_f. \]

The representation of the energy part: Kardamakis proposed that an increasing amount of effort is required for agonist extraocular muscle to contract progressively when eyes move into eccentric positions [4]. So, we consider the energy part as 
\[ J = \int_0^T \alpha(x_c(t))u(t)^2 dt. \]

Where \( \alpha(x_c) = \alpha_0 + \alpha_1 x_c + \alpha_2 x_c^2 \) is polynomial simulation of the relationship between eye position and extraocular muscle tension.

2.1. Eye-head Coordinated Optimal Control Equations

Cost function is as follows:

\[ J = \min \left\{ \int_{t_0}^{t_f} \left[ K + \left[ \alpha(x_e)u_{e}^2 + \beta u_{h}^2 \right] \right] dt \right\} \]

(1)

Where \( t_0 \) and \( t_f \) are start and end time of movements, \( K \) is weighting factor attached to time and energy, \( a_0 = 9.1, a_1 = 0.36, a_2 = 9.1 \) in \( \alpha(x_c) \), \( \beta \) is the weight that scales the head command. \( u_e \) and \( u_h \) are eye and head control signals.

The dynamic characteristics of the eye and head can be described by a second-order system. The equation of state for the coordinated process of the eye and head can be briefly described [4][8] as:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2 \\
\dot{x}_3 \\
\dot{x}_4
\end{bmatrix} =
\begin{bmatrix}
a_{11} & a_{12} & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & a_{33} & a_{34} \\
0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4
\end{bmatrix} +
\begin{bmatrix}
b_1 \\
0 \\
0 \\
0
\end{bmatrix}u_e +
\begin{bmatrix}
0 \\
0 \\
b_3 \\
0
\end{bmatrix}u_h
\]

(2)

\[ y = x_2 + x_4 \]

Where \( a_{11} = -90, a_{12} = -90, b_1 = 4, a_{33} = -15, a_{34} = -6.54, b_3 = -4, x_1, x_2, x_3 \) and \( x_4 \) represent eye’s velocity and position, head’s velocity and position, \( y \) represents angle of gaze shifts.

Using Lagrange multiplier method to formulate the Hamilton function of this paper is as follows:

\[ H(x,u,t) = K + \alpha u_e^2 + \beta u_h^2 + p_1 (a_{11} x_1 + a_{12} x_2 + b_1 u_e) + \\
p_2 x_1 + p_3 (a_{33} x_3 + a_{34} x_4 + b_3 u_h) + p_4 x_3 \]

(3)

Necessary conditions for finding minimum value of cost function derived from Euler equation is equation (4). Together with equation (2), we can get the optimal control differential equations to be solved in this paper. Boundary conditions are \( x(t_0) = [0 \ 0 \ 0 \ 0]^T, x(t_f) = [0 \ \theta_e \ 0 \ \theta_h]^T \), which mean the start velocity and position of eye or head are zero, the end velocity of eye or head is zero, the end
velocity of eye or head is zero, the end position of eye is $\theta_e$ and the end position of head is $\theta_h$. The optimal control signals $u_e$ and $u_h$ can be derived by calculating the function $\frac{\partial H}{\partial u} = 0$.

$$
\begin{align*}
\dot{p}_1 &= -a_{11}p_1 - p_2 \\
\dot{p}_2 &= -a_{12}p_1 - \frac{d(\alpha(x_2))}{dx_2} u_e^2 \\
\dot{p}_3 &= -a_{33}p_3 - p_4 \\
\dot{p}_4 &= -a_{34}p_3
\end{align*}
$$

(4)

$$
u_e = -\frac{b_1 p_1}{2\alpha(x_2)} , \text{ and } u_h = -\frac{b_3 p_3}{2\beta}
$$

(5)

Because of that this paper optimizes time and the terminal time is unknown, we increase the unknown variable $T$ and we have to add another constraint condition to solve differential equations:

$$H(x(t_f), u(t_f), t_f) = 0
$$

(6)

The bvp4c function in MATLAB can be used to solve two-point boundary value problems. However, integral terminal $T$ in differential equations is unknown. So, we do following transformation:

$$
\tau = \frac{t}{T}
$$

(7)

Then integral of $t$ can be converted into integral of $\tau$ and $\tau$, as we have known, is in range of [0,1].

2.2. Results and Discussion

By adjusting the value of $K$, we can get different results. When $K$ approaches $+\infty$, it is time-optimal. And when $K$ approaches 0, it is energy-optimal. Normalizing the time part to the energy part, we can get the solution of equations in line with the characteristics of biological data with the value of $K$ in range of [0.3, 3]. In this paper, some calculation results are selected to illustrate. Figure 1 shows velocity curves of eye and head at different angles of gaze shifts with $K = 0.8$.

![Figure 1](image_url)

**Figure 1.** Velocity curves of eye and head coordinated time-energy optimal control. (a) Results of eye’s velocity curves at different angles of gaze shifts with $K = 0.8$. (b) Results of head’s velocity curves at different angles of gaze shifts with $K = 0.8$.

Figure 1 shows that the velocity curves of eye changes from single peak to double peak with the increase of gaze shifts’ angles, which is consistent with biological data [9]. The velocity curve of head shows a clear parabolic law, which is also the reason why the velocity curve of eye is bimodal. Other
calculation results are shown in Table 1.

**Table 1.** Optimal control solution results with different weights.

| K | variable | 25° | 35° | 45° | 55° | 65° |
|---|----------|-----|-----|-----|-----|-----|
| 0.8 | Time (ms) | 96.9 | 131.6 | 157.7 | 184.1 | 202.6 |
| | Peak velocity (deg·s⁻¹) | Eye | 295.8 | 238.7 | 219.4 | 201.1 | 200.6 |
| | | Head | 66.5 | 122.2 | 156.8 | 193.6 | 217.5 |
| 1.2 | Time (ms) | 85.2 | 117.4 | 141.3 | 165.8 | 182.9 |
| | Peak velocity (deg·s⁻¹) | Eye | 345.6 | 275.8 | 253.1 | 232.9 | 228.8 |
| | | Head | 75.8 | 137.3 | 175.4 | 215.8 | 241.8 |
| 1.6 | Time (ms) | 77.9 | 108.3 | 130.8 | 153.8 | 170.0 |
| | Peak velocity (deg·s⁻¹) | Eye | 384.0 | 304.6 | 280.1 | 258.1 | 254.1 |
| | | Head | 82.9 | 149.1 | 189.8 | 233.1 | 258.2 |

From Table 1 we can find, with the increase of K, time proportion also increases and the result is that the peak velocity of eye and head increases, the terminal time decreases. Besides, the changing process of eye velocity curves from single peak to double peak is slowed down. When K is invariable, peak velocity of eye decreases sharply first, then slowly and peak velocity of head continues to rise at different gaze shifts’ angles. The results are consistent with the trend shown in Figure 1.

### 3. Engineering Model

In order to apply the optimal control signals obtained from section 2 to the practical engineering, this paper considers the above calculation results as the criteria for the selection of hardware components. Finally, the bionic binocular vision system we build is shown in Figure 2.

![Bionic binocular vision system](image)

**Figure 2.** Bionic binocular vision system.

#### 3.1. Eye and Head Velocity Curves Planning

We choose step motors as drive section and gear pair as transmission. Assuming that all the components are rigidly connected, the equations of equilibrium and kinematic can be describe as:

\[
T_M = J\ddot{\theta} + D\dot{\theta} + F\theta + T_l
\]
Where $T_m$ is electromagnetic torque, $T_i$ is constant resisting moment, $J$ is system moment of inertia, $D$ is viscous friction coefficient, $F$ is a factor related to rotation angle, $\theta$ is motor output angle.

According to the above relations, the state equation of the actual system can be described as:

$$
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2 \\
\dot{x}_3 \\
\dot{x}_4 
\end{bmatrix} = \begin{bmatrix}
-D_e J_e^{-1} & -F_e J_e^{-1} & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & -D_h J_h^{-1} & -F_h J_h^{-1} \\
0 & 0 & 1 & 0 
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 
\end{bmatrix} + \begin{bmatrix}
1 \\
0 \\
0 \\
0 
\end{bmatrix} \begin{bmatrix}
\frac{T_{Me} - T_{Fe}}{J_e} \\
\frac{T_{Mb} - T_{th}}{J_h} 
\end{bmatrix}
$$

(9)

$$
y = x_2 + x_4
$$

Where $x_1$, $x_2$, $x_3$ and $x_4$ represent eye’s velocity and position, head’s velocity and position, $y$ represents angle of gaze shifts.

The total energy consumption of motor in the process of robot motion can be divided into two parts. The first part is motion work of driving robot which is integral of joint torque on rotation angle in motion direction. And the second part is energy consumption of heating. Research [13] proposed a brief description of energy consumption as equation (10). Although it cannot be used to directly solve the total energy consumption of robots, it can reflect the impact of motion planning on energy consumption.

$$
E = \int_{t_0}^{t_f} T^2 \, dt
$$

(10)

According to time-energy optimal strategy, the cost function can be described as:

$$
J = \int_{t_0}^{t_f} [k + (T_e^2 + T_h^2)] \, dt
$$

(11)

Where $k$ is weighting factor, $t_i$ is gaze shifts duration, $T_e$ and $T_h$ are eye and head joint torque.

As the biological model data shown in section 2, the velocity of head presents obvious parabolic acceleration and deceleration law during gaze shifts. Due to the influence of head signal, the velocity of eye changes from single-peak parabola to double-peak parabola. In order to facilitate engineering implementation, this paper uses polynomial function to simulate the velocity curve of eye and head. Setting the start velocity and position as 0, we can get following constraint conditions:

$$
\omega_e = at^2 + bt, \quad \omega_h = ct^2 + dt \quad \text{and} \quad \omega_g = \omega_e + \omega_h
$$

(12)

Where $\omega_e$ is the velocity of eye, $\omega_h$ is the velocity of head, $\omega_g$ is the velocity of gaze displacement, $a$, $b$, $c$ and $d$ are optimized parameters to be determined.

Using the uniform speed instead of the change from single peak to double peak, the velocity of eye can be described as:

$$
\begin{align*}
\omega_e &= at^2 + bt & \quad t \leq t_1 \\
\omega_e &= -b^2 (4a)^{-1} & \quad t_1 \leq t \leq t_1 + t_2 \\
\omega_e &= a(t - t_2)^2 + b(t - t_2) & \quad t_1 + t_2 \leq t \leq 2t_1 + t_2
\end{align*}
$$

(13)

Finally, the equality constraints and inequality constraints are given as follows. Where $t_1$ is acceleration duration of eye, $t_2$ is uniform motion duration of eye, $t_i$ is the total time of gaze shifts, $\omega_{max}$ is the peak velocity of eye, $\theta_i$ and $\theta_h$ are motion angles of eye and head, $E_1$ is energy consumption in acceleration and deceleration of eye, $E_2$ is energy consumption in uniform motion of eye, $E_3$ is energy consumption in acceleration and deceleration of head. $\beta_1$ and $\beta_2$ are torque safety factors. $J_e$ and $J_h$ are rotational inertia of eye and head.
\[ \begin{align*}
& t_1 = -\frac{b}{2a} \\
& \omega_{\text{max}} = -\frac{b^2}{4a} \\
& 2 \times \left( \frac{1}{3} a t_1^3 + \frac{1}{2} b t_1^2 \right) + \omega_{\text{max}} t_2 = \theta_e \\
& t_f = 2 \times t_1 + t_2 \\
& -\frac{d}{2} = \frac{t_f}{2} \\
& \frac{1}{3} c t_f^3 + \frac{1}{2} d t_f^2 = \theta_h
\end{align*} \]

\[ \begin{align*}
& t_1, t_2 \geq 0 \\
& a, c \leq 0 \quad b, d \geq 0 \\
& E_1, E_2, E_3 \geq 0 \\
& \max \{ J_e (2a + b) \} \leq \beta_1 T_e \\
& \max \{ J_h (2c + d) \} \leq \beta_2 T_h
\end{align*} \]

(14)

3.2. Results and Discussion

As cost function, inequality constraint conditions and equality constraint conditions have been known, we use genetic algorithm function of MATLAB to solve the problem. The detailed parameters set are shown in Table 2.

| Parameters                        | Value |
|-----------------------------------|-------|
| Maximum number of generations    | 100   |
| Population size                  | 60    |
| Gene length                       | 20    |
| Generation gap                    | 0.9   |
| Crossover rate                    | 0.7   |
| Mutation rate                     | 0.01  |

The calculated results are in accordance with biological characteristics when \( k \) is in range of [0.15, 0.5]. So, a numerical example with the gaze shifts’ angle at 75° is given to illustrate the validity of our method. And the value of \( k \) is 0.3, the value of \( K \) is 0.8. The results are shown in figure 3.

![Figure 3](image-url)  
**Figure 3.** Gaze shifts time-energy optimal control velocity curve with the angle of 75º.  
(a) Velocity curve of biological model with \( K = 0.8 \).  
(b) Velocity curve of Engineering model with \( k = 0.3 \).
Table 3 shows the detailed results. As can be seen from the table 3 and figure 3, the results of section 3 the actual engineering system planning is similar to the results obtained from section 2 biological data analysis. The process from single peak to double peak of eye velocity has no practical significance for engineering so that we use uniform motion to instead it. The velocity of head takes on obvious parabolic acceleration and deceleration law. To avoid impact or impulse affecting image quality of cameras, we add a constraint condition that acceleration decreases to 0 at the final stage of acceleration and acceleration is also 0 at the beginning of deceleration stage. Meanwhile, the peak velocity is also limited by this condition. With the same gaze shifts angle, the results of engineering system planning method are similar to the results of biological model on the value of motion duration and peak velocity, which demonstrates that our bionic vision system is quite qualified for the simulation of primates’ coordinated eye and head movements to complete gaze shifts.

**Table 3.** Detailed results of biological model and engineering model.

| Weight | Times(ms) | Position(deg) | Peak velocity(deg·s⁻¹) |
|--------|-----------|---------------|------------------------|
| t₁     | t₂        | eye           | head                   | eye | head |
| 0.8    | 221.8     | 42.29         | 36.48                  | 204.5 | 238.2 |
| 0.3    | 85.5      | 58.1          | 42.29                  | 36.48 | 249.8 | 245.8 |

4. Conclusion
This paper analyses the mechanism of coordinated eye and head movements of primates and builds bionic orchard robot vision system. With the purpose of time-energy optimal control, this paper plans eye and head velocity curves during gaze shifts. And the results are similar to the biological data, which verifies the feasibility and validity of this method. In addition, the method proposed in this paper is not limited to the structure of figure 2, but provides a general solution for eye and head coordinated movements of orchard robot vision system. Future work is to realize automatic searching and tracking of stationary or moving targets with two eyes and head coordinated movements.

5. Acknowledgments
This work is supported by Key Research and Development Program of Shaanxi Province (2018NY-111).

6. References
[1] Li Shuping, Cheng Jun and Li Hengyu 2016 *Computer Engineering* 42(9): 273-78.
[2] Mao Xiaobo 2011 *Journal of Biomedical Engineering* 28(5): 895-900.
[3] Wang Xiaohui, Ma Xin and Qin Bin 2016 *Journal of System Simulation* 28(6): 1380-85.
[4] Kardamakis AA, Moschovakis AK 2009 *The Journal of Neuroscience* 29(24): 7723-30.
[5] Sohrab Saeb, Cornelius Weber and Jochen Triesch 2009 *International Joint Conference on Neural Networks* (Atlanta/USA)
[6] Saglam M, Lehnen N and Glasauer S 2011 *Journal of Neuroscience* 31(45): 16185-93.
[7] Saeb S, Weber C, Triesch J and Diedrichsen J 2011 *PLoS Computational Biology* 7(11):e1002253.
[8] Wang Xuezong, Simon MH 2011 *Biological Cybernetics* 104(1/2): 65-73.
[9] Freedman EG 2001 *Biological Cybernetics* 84(6): 453-62.
[10] Freedman EG 2008 *Experimental Brain Research* 190(4): 369-87.
[11] Moschovakis AK 2010 *Biological Cybernetics* 102(3): 209-25.
[12] Nelson WL 1983 *Biological Cybernetics* 46(2): 135-47.
[13] Gregory J, Olivares A and Staffetti E 2012 *Systems & Control Letters* 61(2): 279-91.