Modeling and Dynamic Analysis of Three-Degree-of-Freedom Spherical Actuator under Deep Reinforcement Learning Control

Hirotugu Fusayasu¹, Akira Heya², and Katsuhiro Hirata²

¹Panasonic Corporation, Kadomashi 571-8506, Japan, fusayasu.hirotugu@jp.panasonic.com
²Osaka University, Osaka 565-0871, Japan, akira.heyu@mapse.ing.osaka-u.ac.jp

Multi-degree-of-freedom (multi-DOF) spherical actuators have been developed for the fields of robotics and industrial machinery. We have proposed an outer rotor type three-DOF spherical actuator that can realize a high torque density. Each coil input current is calculated using a torque generating equation based on the torque constant matrix. Permanent magnet type actuators have a problem that a PID control limit is exceeded when a cogging torque which is different from design is generated, which is a problem due to manufacturing fluctuation factors. Therefore, we applied a deep neural network that is expected to realize a strong non-linearity control of spherical actuators, and introduced a feedforward current compensator by reinforcement learning, and applied it to uncertainty problems such as manufacturing fluctuations of cogging torque.

Index Terms—multi-degree-of-freedom, spherical actuator, deep reinforcement learning, industrial robots, torque constant map.

I. INTRODUCTION

MULTI-DOF spherical actuator achieves high-speed and high-level operation by downsizing the mechanism and simplifying control and is expected to be applied to the field of industrial robots and to replace human workers, especially for detailed and dangerous tasks. We previously proposed a method to calculate the input current value using a torque generation formula based on the torque constant map [1]. We have shown multi-DOF spherical actuator applicability to a precision robot and evaluated dynamic torque characteristic analysis accuracy [2].

It has been reported that permanent magnet motors generate unexpected cogging torque due to various manufacturing errors [3]. Manufacturing errors mean between the ideal condition at the motor design phase and the actual condition in mass production. A cogging torque may be about 2 to 3 times that of the ideal time, and the same problem is assumed in the spherical actuator. In this case, the actuator PID control limit would be exceeded. On the other hand, some studies introduced neural networks and fuzzy controls for the disturbance of uncertainty of spherical actuators [4]-[5]. However, they are difficult to deal with the case where the magnitude of the disturbance is large or the nonlinearity is strong.

In this paper, we propose a method control which is suitable for spherical actuators using a deep neural network that realizes a strong non-linearity control and a feedforward current compensator, and applies it to uncertainty problems such as manufacturing variation of cogging torque.

II. ACTUATOR STRUCTURE AND OPERATION PRINCIPLE

An analyzed outer rotor spherical actuator is shown in Fig. 1. 10 coils are arranged in the stator, and 5-phase currents are applied to rotate the rotor. A torque equation is given as follows. This assumes that each coil magnetic field does not interfere with the other coil magnetic fields and independently contributes to the torque of the actuator.

\[ T = K_m i + T_{cog} \]  

(1)

where \( T \) is an output torque, \( K_m \) is a magnet torque constant, \( i \) is a coil current of each phase, \( T_{cog} \) is a cogging torque, respectively. As shown in (1), the spherical actuator torque is obtained by summing up the torques generated by a simple magnetic pole model shown in Fig. 1(d), which consists of a pair of magnetic poles at each position. A magnetic field analysis using a 3D FEM is performed for each posture of the simple magnetic pole model, and the magnet torque constant matrix \( K_m \) and the cogging torque vector \( T_{cog} \) are calculated.

These torque constant matrices are maps that describe the torque constant around the \( x-, y-, \) and \( z- \) axes concerning present magnetic pole positions. The torque constant required at the magnetic pole position for any rotor posture is obtained by referring to the value corresponding to the magnetic pole position from the torque constant map.

III. DYNAMIC ANALYSIS

Fig. 2 shows the closed-loop control system for the dynamic torque characteristic analysis. In the current control, the phase currents of the order to rotate the target angle at any position is calculated by multiplying both sides of (1) by the inverse matrix of \( K_m \). In the torque analysis, the torque acting on the rotor is calculated by the current of each phase, and the analyzed angle

![Fig. 1. Basic structure of outer rotor spherical actuator.](Image 1)

![Fig. 2. System configuration of dynamic analysis.](Image 2)
is calculated by solving the equation of motion (EOM). We performed a torque error tolerance analysis. A target angle was set to a sine wave with an amplitude of 15° and 2 Hz around the x-axis, with an amplitude of 15° and 1 Hz around the y-axis and with an amplitude of 10° and 1 Hz around the z-axis. Assuming a manufacturing error, the cogging torque was increased to 3 times the designed value, and the influence on the dynamic analysis was investigated. The root mean square error (RMSE) was shown in Fig. 3. From Fig. 3, it can be seen the RMSE increases as the cogging torque increases. Fig. 4 shows the analysis results of the rotation angle when cogging torque was twice the design value. The dashed line represents the target angle, and the solid line represents the analyzed angle. From Fig. 4, it can be seen that the vibration component increases and it becomes difficult to follow the rotation angle with PID control. Figs. 3 and 4 show that the RMSE was 0.7° and the maximum error was 1.8°.

IV. INTRODUCTION OF DEEP REINFORCEMENT LEARNING

We present a feedforward current compensator using deep reinforcement learning. Fig. 5 shows a closed-loop control system for dynamic torque characteristic analysis with a feedforward current compensator. Reinforcement learning is an algorithm that imitates the psychology of learning of human beings, and is a method of sequentially searching from the environment in the order of state - action - reward - action - state - action. In this investigation, we used deep deterministic policy gradient (DDPG) [6]. This is a kind of reinforcement learning algorithm, which can be used when the behavioral space is continuous. DDPG trains two action evaluators called actors and critics at the same time to learn the best policies to maximize long-term rewards. Both the actor representing the policy and the critic representing the value function are represented by a neural network. The critic parameters are updated to minimize the loss function $L$.

$$y_t = r_t + \gamma Q'(s_{t+1}, \mu'(s_{t+1})|\theta^\phi)$$

$$L = \frac{1}{N} \sum_i (y_t - Q(s_i, a_i|\theta^\phi))^2$$

where $Q(s, a)$ is a state action value function, $\tau$ is a time step of the reinforcement learning agent, $s_t$ is a state, $a_t$ is an action, $r_t$ is a reward to be acquired, $N$ is number of mini-batch. The actor parameters are updated using the gradient below.

$$\nabla_{\theta^\mu} L = \frac{1}{N} \sum_i \nabla_{\theta^\mu} Q(s, a|\theta^\phi)_{s=s_i, a=a_i} \nabla_{\theta^\mu} \mu(s|\theta^\phi)_{a_i}$$

where $\mu$ is an actor function, $\theta^\phi$ and $\theta^\phi$ are the parameters of the current critic and actor network, respectively, $\theta^\phi$ and $\theta^\phi$ are the parameters of the target critic and actor network, respectively, and $\gamma$ is the discount rate. In the current compensator, the rotation angle error, the rotation angle, and the angular velocity are input to the state $s_t$, and the compensation current of each phase is output to the action $a_t$.

**REFERENCES**


