Truss Assembly by Space Robot and Task Error Recovery via Reinforcement Learning

Proceedings of the 2000 IEEE/RSJ
International Conference on
Intelligent Robots and Systems

Truss Assembly by Space Robot and Task Error Recovery via Reinforcement Learning

Kei Senda

Graduate School of Engineering
Osaka Prefecture University
1-1 Gakuen-cho, Sakai, Osaka 599-8531, Japan
senda@aero.osakafu-u.ac.jp

Tsutomu Matsumoto

Graduate School of Engineering
Osaka Prefecture University

Abstract

This paper addresses an experimental system simulating a free-flying space robot, which has been constructed to study autonomous space robots. The experimental system consists of a space robot model, a frictionless table system, a computer system, and a vision sensor system. The robot model composed of two manipulators and a satellite vehicle can move freely on a two-dimensional planar table without friction by using air-bearings. The robot model has successfully performed the automatic truss structure construction including many jobs, e.g., manipulator berthing, component manipulation, arm trajectory control avoiding collision, assembly considering contact with the environment, etc. Moreover, even if the robot fails in a task planned in advance, the robot accomplishes it by task re-planning through reinforcement learning. The experiment demonstrates the possibility of the automatic construction and the usefulness of space robots.

1 Introduction

Space robots are necessary for future space projects to construct, repair and maintain satellites and space structures in orbits. A space robot consists of manipulators and a satellite base, which can fly freely in an orbit. This type of space robots is called free-flying space robot (this paper calls it just a space robot).

Lots of new complicated dynamic problems have been raised, e.g., an interaction between the manipulators and satellite, a structural flexibility caused by lightweight requirements, etc. There exist many researches focused on the dynamic problems [1]–[5]. Some studies using hardware equipments on the ground have been reported to examine the control and identification methods [5]–[12]. Moreover, studies of autonomous systems, e.g., utilization of the force and vision information, planning and reasoning, etc., are necessary to realize the autonomous space robots that can achieve their mission commanded by human operators [13]. There have been some projects emphasize the present point [11]–[16]. Because all of them are huge projects, simple testbeds on the ground are requested for the research and development of autonomous space robots.

Hence, this study has developed a ground experimental system simulating a free-flying space robot under micro-gravity condition in orbit (**Fig. 1**). In this study, the space robot model assembles a truss structure automatically. Repeating the sequence enables us to construct large structures.

But, the space robot may fail in a task planned in advance because of uncertainties and variations of the work site. To accomplish the task, the robot must modify the task-sequence suitable for the real environment. For the purpose, the robot re-plans the task-sequence by using reinforcement learning. It then achieves the new task-sequence experimentally.

The rest of this paper is organized as follows. The experimental system is introduced in the next section. The third section explains fundamental techniques realizing the assembly, e.g., the stereo image measurement, the visual servoing, the positioning control of free-floating space robot, path planning of arms avoiding collision with the environment, the force controls considering contact with the environment, etc. In section 4, the automatic truss structure assembly is experimentally demonstrated by synthesizing the tech-

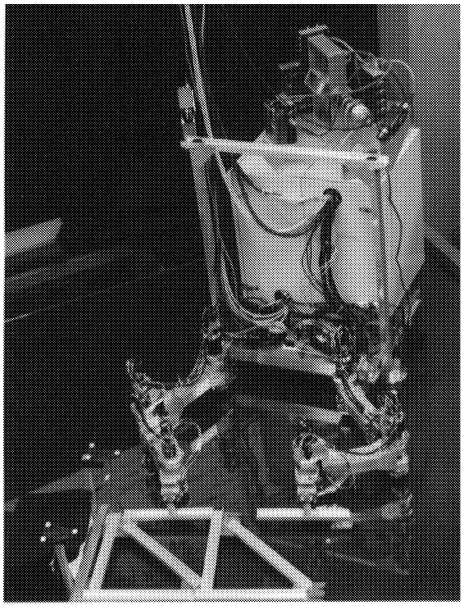

Figure 1: Photograph of space robot model and truss

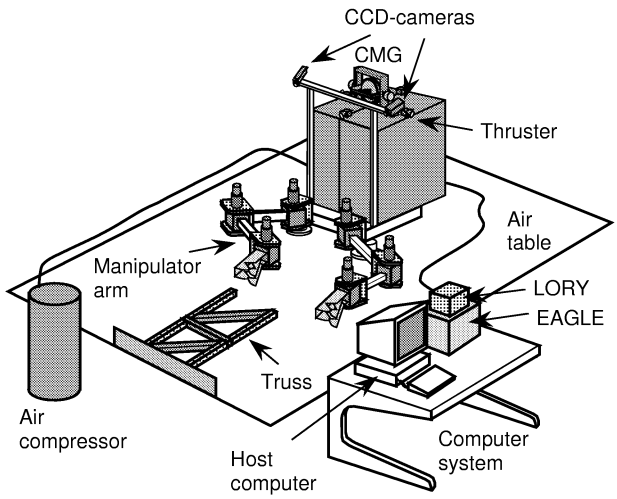

Figure 2: Schematic diagram of experimental system

niques. Section 5 illustrates the method using reinforcement learning to plan task-sequence appropriately for the real environment when the robot fails in the task planned in advance. Some concluding remarks are given in the final section.

2 Experimental System

2.1 Outline of Experimental System

Fig. 1 is a photograph of the space robot model and a truss structure under assembling. **Fig. 2** shows a schematic diagram of the experimental system constructed in this study. The experimental system simulates a free-flying space robot in orbit while motion of the robot model is restricted in a two-dimensional plane. The experimental system consists of four subsystems: (a) space robot model, (b) frictionless table system, (c) computer and I/O system, and (d) vision sensor system. The robot model is supported on the horizontal table without frictions by using air-pads. Its micro-gravity condition was assessed by the average translation acceleration during motion, which was estimated below $1.0 \times 10^{-4}g$ where g is the unit of the gravitational acceleration [10]. For reference, the average acceleration of the orbital space shuttle is approximately $1.0 \times 10^{-4}g$.

Information from the robot model is put into the computer system placed beside the table. In the vision sensor system, the stereo images are taken by the CCD cameras and sent to an image-processing unit. After appropriate process in the image-processing unit, the visual information is sent to the computer system. The computer system processes the sensing data and computes control commands to the robot model. See reference [17] for details of the experimental system.

2.2 Specifications of Experimental System

(a) Space Robot Model The robot model consists of a satellite vehicle and dual 3-DOF SCARA type rigid manipulators. A position/attitude control system and the CCD cameras for a stereovision are installed on the satellite vehicle. Applied forces and torque at the end-effector can be measured. The hand has one DOF of open/close states, and can grasp a payload. The parameters of the robot model are listed in **Table 1**, where a_i and i_i indicate the position of mass center and the mass moment of inertia of link i about its mass center, respectively. The links and joints are numbered from the base to the tip of the manipulators, where link 0 and link 3, respectively, are the satellite vehicle and a manipulator hand. The both manipulators have the same specifications. All joints are driven by DC servomotors with "harmonic drive" reduction gears. The satellite vehicle has a position/attitude control system with thrusters and a control momentum gyro [18].

(b) Frictionless Table System The frictionless table system realizes a two-dimensional weightless condition in the horizontal plane. The table system consists of three parts: the horizontal planar table, air-pads, and a pressure air source. The top of the table

Table 1: Specification of space robot model

Link i	Mass m_i[kg]	Length ℓ_i[m]	Mass center a_i[m]	Inertia i_i[kg·m^2]
Link 0	66.36	0.464	0.000	5.917
Link 1	2.96	0.320	0.182	0.045
Link 2	2.25	0.260	0.132	0.028
Link 3	1.70	0.118	0.031	0.002
Payload	1.95		0.000	0.020

is leveled carefully to realize a horizontal flat plane on the tabletop. Air-pads are used to support the space robot model on the planar table without friction.

(c) Computer and I/O System The computer and I/O system is composed of three sub-systems: a host computer, EAGLE-10 with a CPU and an image processing board, and LORY with several boards for the robot control.

(d) Vision Sensor System The vision sensor system has two CCD cameras for a stereo image to measure the position and orientation of target objects. Details of the stereo measurement are explained later.

3 Subsystems for Control

(a) Stereo Vision Measurement The spatial positions of reference points in the environment are calculated from the stereo camera image by using the triangulation. A target marker with three light-emitting-diodes (LEDs) is used to measure the position and orientation of the target, where the LEDs are arranged in vertexes of an isosceles triangle. By using the stereovision system, the three-dimensional positions and the spatial orientation of the target marker can be measured every 33.3 ms. A time delay of the position and orientation measurement is 66.7 ms. After a hand-eye calibration, the stereovision using the obtained parameters yields the spatial position measurement with 2 mm mean error.

(b) Visual Servoing The space robot must measure the relative position and orientation to the environment or the manipulated objects by using a visual sensor and control its manipulators because the position and attitude of the satellite vehicle changes during tasks. To achieve such a control, a visual servoing has been developed by integrating the information from a vision sensor into the manipulator control loop. In this study, some control systems have been examined by replacing the arm controller. In the visual servoing, the sampling time of the visual sensor is 33.3 ms and that of the control loop is 1 ms.

(c) Positioning Control of Space Robots Control methods for free-floating space robots should be used for some tasks, e.g., manipulator berthing, because manipulator motions vary the position/attitude of the satellite vehicle. The sensory feedback control for space robots [2] is employed in this study that is induced by the following conic artificial potential:

$$\boldsymbol{\tau} = -\boldsymbol{J}^T(\theta_0,\ \boldsymbol{q})\boldsymbol{u} - \boldsymbol{K}_D\dot{\boldsymbol{q}} \quad (1)$$

$$\boldsymbol{u} = -\frac{\partial \boldsymbol{U}(\boldsymbol{y}_e)}{\partial \boldsymbol{y}_e} \quad (2)$$

where $\boldsymbol{y}_e = \boldsymbol{y} - \boldsymbol{y}_d$, $\boldsymbol{y}$ the position/orientation of hand, $\boldsymbol{y}_d$ the desired position/orientation, $\boldsymbol{q} = [\theta_1, \ldots, \theta_n]^T$ the joint angles, $\boldsymbol{J}(\theta_0,\ \boldsymbol{q})$ the generalized Jacobian matrix [1], θ_0 the vehicle attitude angle, $\boldsymbol{K}_D$ the velocity feedback gain matrix, $\boldsymbol{U}(\boldsymbol{y}_e)$ the conic artificial potential, and $\boldsymbol{\tau}$ the joint input torque. The control scheme makes the joint angular velocities and the position/orientation error of the manipulator hands converge to zeros, and the asymptotic stability of a static target is guaranteed. Further, the control scheme is compatible with the visual servoing because it enables us to feedback the manipulation variable error measured in the satellite frame.

(d) Path Planning and Trajectory Control of Arms After the robot berthed to the work site, it manipulates components to be assembled, plans the path of arms avoiding collision with the environment, and controls the arms to track the obtained path. Here the path is planned to avoid obstacles using an artificial potential method proposed by Tsuji et al.[20], where objects in the environment, e.g., truss under assembly, are regarded as obstacles.

(e) Force Control Contacting with Environment For the assembly, the space robot applies some force to the truss, e.g., pushing a truss component into a connector, and receives its reaction. Consequently, the space robot works after fixing the satellite vehicle to the task site. In the situation, the space robot is similar to the base-fixed robot on the ground. On the other hand, force control is needed considering constraints due to contact with the environment. This study uses the position-force hybrid control called SP-DF control as follows [21]. The equation of motion of the robotic manipulator constrained at the hand is

$$\boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}} + \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = (\nabla_q \boldsymbol{\phi})^T \frac{\boldsymbol{f}}{||\nabla_y \boldsymbol{\phi}||} + \boldsymbol{\tau} \quad (3)$$

where $\boldsymbol{M}$ denotes the inertia matrix, $\boldsymbol{h}$ the centrifugal and Coriolis forces, $\boldsymbol{f}$ the external forces applied to the

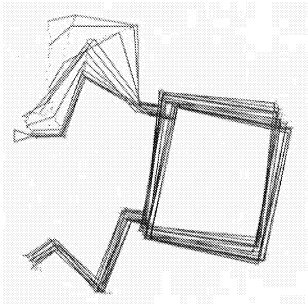

Figure 3: Experimental result of manipulator berthing

hand, $\boldsymbol{\phi}(\boldsymbol{q}) = \mathbf{0}$ or $\boldsymbol{\phi}(\boldsymbol{y}) = \mathbf{0}$ the constraint equation, and the first term in the right-hand side is the reaction transformed into the joint space. The control input $\boldsymbol{\tau}$ is given as

$$\boldsymbol{\tau} = -B_0\dot{\boldsymbol{q}} - A_0\mathrm{Sin}(\Delta\boldsymbol{q}) - (\nabla_q\boldsymbol{\phi})^T\frac{\boldsymbol{f}_d}{\|\nabla_y\boldsymbol{\phi}\|} \qquad (4)$$

where $\boldsymbol{f}_d$ is the desired $\boldsymbol{f}$ and $\mathrm{Sin}(\theta) = -\gamma\sin(\gamma\theta)$ saturating at $\pi/2\gamma \leq \|\theta\|$. The $\nabla_q\boldsymbol{\phi}$ can be obtained from the contact reaction. The control scheme guarantees that $\dot{\boldsymbol{q}}(t) \to 0$ as $t \to \infty$ if $\boldsymbol{q}_d \equiv$ constant, $\boldsymbol{\phi}(\boldsymbol{q}_d) = 0$ holds, and the hand is to move on $\boldsymbol{\phi}(\boldsymbol{q}) = 0$.

4 Truss Assembly Experiment

The robot performs several tasks in the truss assembly, e.g., the manipulator berthing, the component manipulation, the arm trajectory control avoiding collision, the assembly considering contact with the environment, etc. This section illustrates parts of them.

An experimental result of the manipulator berthing is shown in **Fig. 3**, where the visual servoing with the sensory feedback control for space robots is used. The right manipulator hand is controlled well and the manipulator berthing is successful, whereas the satellite vehicle is moved by the reaction of the arm motion and the disturbance of cables suspended from above.

Fig. 4 shows an experimental result of truss assembly. In the sequence, manipulating a truss component and connecting it to a node proceed the assembly process. The component is installed in the planned position and direction because the connector at the node has a notch to insert the component. The installed component would not be detached since the connector has a latch mechanism. By using above method, the arm path is planned and the manipulator is controlled to track the obtained path. The component installation is performed well by the SP-DF control. Fig. 1 corresponds to Step 3.

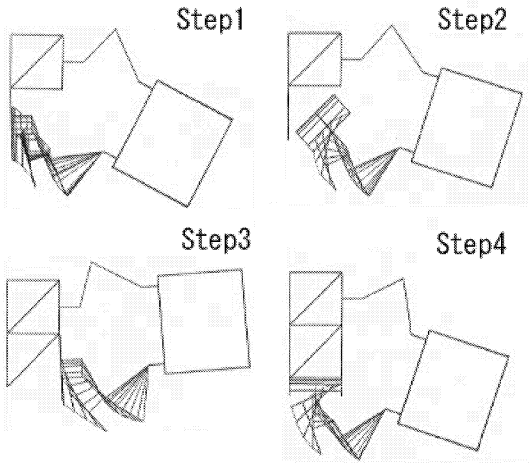

Figure 4: Experimental result of truss assembly

5 Task Error Recovery by Learning

In the previous section, the robot has successfully achieved the truss structure assembly by the task-sequence planned in advance. However, the space robot may sometimes fail in the task by the sequence because of uncertainties and variations of the work site. To recover from the error and accomplish the task, the robot must re-plan the task-sequence suitable for the real environment. For the purpose, reinforcement learning with trial and error is applied.

5.1 Reinforcement Learning

For the re-planning, one of typical reinforcement learning algorithm, Q-learning [22], is used. The reinforcement learning is used because the robot learns how to do suitably for the real environment so as to maximize a numerical reward signal r that is given by the designer to describe what to do.

The Q-learning estimates the optimal action-value function $Q(s, a)$, which gives the best action a for the state s, through interactions between the learner and the environment with trial-and-error processes. It has been shown that the estimated action-value function Q converges with probability one to the optimal if the problem is a finite Markovian decision process. This study uses the following algorithm.

Initialize $Q(s, a)$ arbitrarily
Repeat until learning convergence
  Initialize s
  Repeat until episode terminal
    Choose a from s using policy derived from Q
    Take action a and observe new state s' and r

Update Q by

$$(1-\alpha)Q(s,a) + \alpha[r + \gamma \max_{a'} Q(s',a')] \quad (5)$$

Go next step if s is terminal

Stop if learning is convergent

In Eq. (5), the learning rate is α $(0 < \alpha \leq 1)$ and the discount rate is γ $(0 < \gamma \leq 1)$. This study uses ϵ-greedy policy [22] to choose action through learning.

5.2 Example Problem

Consider the situation that the diagonal element is luck after Step 3 during the truss structure assembly sequence of Fig. 4. In this situation, the robot cannot assemble the diagonal element into the truss structure by the sequence planned in advance because the element attached in Step 3 becomes a new obstacle.

It is difficult to plan the path to the goal avoiding obstacles by the artificial potential method [20] when the work site is complicated like this situation. The artificial potential method defines a potential field in the state space, makes the potential minimum at the desired state, and plans the path by using a non-linear programming. This method requires that the defined potential field have only one global minimum to solve the problem. Because the environment is complicated, it is not easy to make the potential field without local minimum. In the meantime, the reinforcement learning can overcome the problem.

5.3 Experimental Result and Discussion

A discrete state space with $15^3 = 3375$ states is made for the reinforcement learning, where each coordinate of hand position $(x,\ y)$ and orientation θ is discretized in 15. The x and y coordinates are quantized every 0.05m and θ every 10deg. Each state has $2 \times 3 = 6$ actions that are one-step movements of discrete coordinates to the neighbor. Parameters in Eq. (5) for updating Q are $\alpha = 0.1$, $\gamma = 0.6$, and $r = 10$. For the ϵ-greedy policy, $\epsilon = 10$ is initially used and reduced gradually to be the policy deterministic.

Fig. 5 is the experimental result where arm path is planned by the reinforcement learning. The generated task-sequence in the state space is illustrated in **Fig. 6**. The initial and the target states of (x, y, θ) are $(0.42, 0.3, -15.0)$ and $(0.62, 0.3, -35.0)$, respectively. The learning method enables to accomplish the task avoiding collision against the environment.

Most path planning methods generate a path from the initial state to the desired state. On the other hand the reinforcement learning estimates the optimal

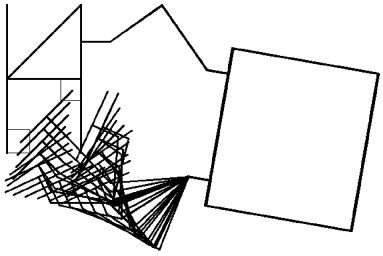

Figure 5: Experimental result of reinforcement learning

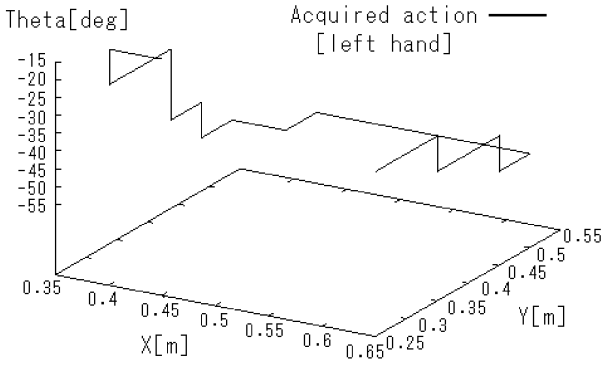

Figure 6: Acquired action of position and orientation of left hand

action-value function Q for all states that derives a policy to choose the best action. Therefore, the robot can make the best action at any state in the environment after the optimal Q is estimated.

6 Concluding Remarks

This study has demonstrated the automatic truss structure assembly by the experimental autonomous space robot system. The fundamental techniques have been developed and synthesized for the assembly, i.e., the stereo image measurement, the visual servoing, the positioning control of free-floating space robot, the arm path planning, and the force control considering contact with the environment. The robot has successfully achieved the automatic truss structure assembly. Furthermore, the robot has re-planned the task-sequence by using reinforcement learning and achieved it even if the robot failed in the task-sequence planned in advance. As a result, this study has shown a possibility of the automatic truss structure construction and the usefulness of space robots.

There remain some subjects for real space robots. The truss assembly has been accomplished by the se-

quential control whereas it is not robust enough against uncertainties in the work site. The intelligence or autonomy is the biggest subject to realize a useful space robot although this study has approached by the reinforcement learning. A real-time stereovision is also a subject, especially it must work in bad lighting environments. Some difficulties may arise to realize robots with 3-dimensional motion, but they will be solved by existing technologies.

Acknowledgement

A part of this work is financially supported by the Special Research Promotion Foundation for Graduate School of Osaka Prefecture University.

References

[1] Umetani, Y. and Yoshida, K., "Continuous Path Control of Space Manipulators Mounted on OMV," *Acta Astronautica*, **15**-12, 981–986, (1987).

[2] Masutani, Y., Miyazaki, F., and Arimoto, S., "Sensory Feedback Control for Space Manipulators," *IEEE Int. Conf. on Robotics and Automation*, Scottsdale, 1346–1351, (1989).

[3] Yamada, K. and Tsuchiya, K., "Efficient Computation Algorithms for Manipulator Control of a Space Robot," *Trans. of Society of Instrument and Control Engineers*, **26**-7, 765–772, (1990). (in Japanese)

[4] Murotsu, Y., Tsujio, S., Senda, K., and Hayashi, M., "Trajectory Control of Flexible Manipulators on a Free-Flying Space Robot," *IEEE Control Systems*, **12**-3, 51–57, (1992).

[5] Murotsu, Y., Senda, K., Ozaki, M., and Tsujio, S., "Parameter Identification of Unknown Object Handled by Free-Flying Space Robot," *AIAA J. of Guidance, Control, and Dynamics*, **17**-3, 488–494, (1994).

[6] Koningstein, R., Ullman, M., and Cannon, Jr., R. H., "Computed Torque Control of a Free-Flying Cooperating-Arm Robot," *NASA Conf. on Space Telerobotics*, Pasadena, (1989).

[7] Yoshida, K., "Experimental Study on the Dynamics and Control of a Space Robot with the EFFORTS Simulators," *Advanced Robotics*, **9**, 583–602, (1995).

[8] Komatsu, T., Ueyama M. and Iikura S, "Autonomous Satellite Robot Testbed," *Proc. of i-SAIRAS*, Kobe, Japan, 113–116, (1990).

[9] Toda, Y., et al., "Development of Free Flying Space Telerobot: Ground Experiments on Two-Dimensional Flat Testbed," *AIAA Guidance, Navigation and Control Conference*, Hilton Head, SC, 33–39, (1992).

[10] Murotsu, Y. et al., "Experimental Studies for Control of Manipulators Mounted on a Free-Flying Space Robot," *Proc. of IROS*, Yokohama, Japan, 2148–2154, (1993).

[11] Kan, E. P., "Supervised Robotics for Space Servicing", *Proc. of IROS*, Tsuchiura, Japan, (1990).

[12] Shimoji, H. et al., "Autonomous Capture Experiment of Free-Flying Target On the Zero Gravity Simulator," *Proc. of i-SAIRAS*, Kobe, Japan, 77–80, (1990).

[13] Skaar, S. B. and Ruoff, C. F. (eds.), "Teleoperation and Robotics in Space," American Institute of Aeronautics and Astronautics, Inc., (1995).

[14] Brunner, B., Hirzinger, G., Landzettle, K., and Heindl, J., "Multisensory Shared Autonomy and Tele-Sensor-Programming — Key Issues in the Space Robot Technology Experiment ROTEX," *Proc. of IROS*, Yokohama, Japan, 2123–2139, (1993).

[15] David, L., "Robots for All Reasons," *Aerospace America*, September, 30–35, (1995).

[16] Special Issue, "Robot Experiments on ETS-VII", *J. Robotics Society of Japan*, **17**-8, 1055–1104, (1999). (in Japanese)

[17] Senda, K. et al., "A Hardware Experiment of Space Truss Assembly by Using Space Robot Simulator," *Proc. of 9th Workshop on Astrodynamics and Flight Mechanics*, Sagamihara, Japan, 233–240, (1999).

[18] Senda, K., Murotsu, Y., Nagaoka, H., and Mitsuya, A. "Attitude Control for Free-Flying Space Robot with CMG," *AIAA Guidance, Navigation, and Control Conf.*, Baltimore, MD, 1494–1502, (1995).

[19] Takegaki, M. and Arimoto, S., "A New Feedback Method for Dynamic Control of Manipulators," *ASME J. Dyn. Sys., Mea., and Con.*, **102**, June, 119–125, (1989).

[20] Tsuji, T. et al., "Motion Planning for Manipulators Using Artificial Potential Field Approach that can Adjust Convergence Time of Generated Arm Trajectory," *J. Robotics Society of Japan*, **13**-2, 285–290, (1995). (in Japanese)

[21] Arimoto, S., "To Make Mechanical Systems More Intelligent III: Quasi-Natural Potential, Hyper-Stable PID Servo, Linear Velocity Observer, and Joint-Space Orthogonalization Principle," *J. Robotics Society of Japan*, **12**-3, 415–425, (1994). (in Japanese)

[22] Sutton, R. S. and Barto, A. G., "Reinforcement learning: An introduction," MIT Press, (1998).