The Fourteenth International Symposium on Artificial Life and Robotics 2009 (AROB 14th '09),
B-Con Plaza, Beppu, Oita, Japan, February 5 - 7, 2009

Control of Damping with Reinforcement Learning for Power-assisted Positioning Task

Tetsuya Morizono[1] and Masatake Higashi[2]

[1]*Fukuoka Institute of Technology, 3-30-1, Wajiro-higashi, Higashi-ku, Fukuoka, 811-0295, Japan*
morizono@fit.ac.jp
[2]*Toyota Technological Institute, Tempaku-ku, Nagoya, Japan*

Abstract: Finding an appropriate reference of kinetic characteristic is a major problem of an impedance-controlled power-assist robot. In this paper, autonomous adjustment of damping (viscosity) based on subjective operational feeling of an operator is discussed. For autonomous adjustment, reinforcement learning is utilized. For adaptation of the robot to a positioning task including multi goal positions, a method for inference of the goal position is developed. Experimental results show that the method developed in this paper is capable of adjusting viscosity of the robot so that dissipation of kinetic energy of the robot assists positioning of an operator at the goal position.

Keywords: Control, Damping, Reinforcement Learning, Power-assist, Positioning, Operational feeling

I. INTRODUCTION

When a power-assist robot is controlled by the impedance control, a major problem is how to find a reference of mechanical impedance appropriate for assistance of an operator. Many studies prefer variable impedance control where control parameters (desired mass, viscous coefficient and/or stiffness) are variable (for example [1]), and operational force and even stiffness of the operator's arm was used as reference signals for adjustment of the control parameters [2][3].

In contrast with those studies, Yamada et al. proposed a scheme for adjustment of control parameters named "Field Impedance Equalizer (FIE)" [4]. The study assumed a repetitive power-assisted positioning task, and the proposed scheme aimed at tune up of the parameters in interaction between an operator and a power-assist robot. Experimental results showed that the viscous coefficient could be adjusted so that an operator obtained a good subjective operational feeling of a robot. However, autonomous adjustment of the coefficient has not been well studied. Therefore, the authors have proposed a method of autonomous adjustment of the viscous coefficient based on FIE [5].

The previous study assumes that the operational distance of a positioning task is constant. However, many tasks existing in factories (an assembly task of automobiles, for example) include different goal positions between operations, and the goal position is often determined by a worker in real time. To extend the previous studies to the case where the operational distance is different between operations and the distance is determined in real time, this paper proposes EFDA (Enhanced Field Damping Adjuster).

II. FIE AND PARAMETER ADJUSTMENT

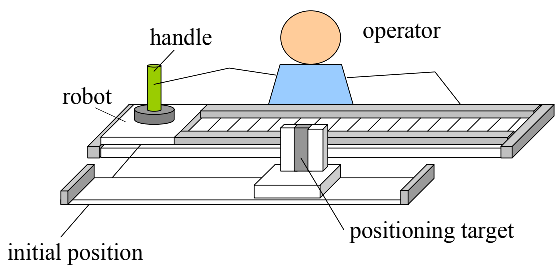

Fig. 1 A 1-DOF power-assist robot

The study in this paper considers a repetitive positioning task with a 1-DOF power-assist robot, as shown in Fig. 1. One task includes repetition of operations from the initial position to the goal position. An operation is one-way; after the robot is positioned at the goal position, return of the robot to the initial position is automatic.

If the goal position is fixed throughout the task, the operating distance is fixed. In this case, the distance can be divided into several sections with constant lengths. The previous study by Yamada et al. [4] defined each of the sections as a "field". *Impedance* parameters of mass, viscous coefficient and stiffness were defined dependent on the field, and tuned (*equalized*) based on operational feeling of the robot. This is the framework of FIE.

Based on FIE, the authors proposed a method of autonomous adjustment. The method adjusted the viscous coefficient, and reinforcement learning was applied for autonomous adjustment. In the method, the state and the action of the agent was defined as the field number and the value of the viscous coefficient, respectively. The action value functions $Q(i, d_i)$ of all the fields were updated by Q-learning, here i and d_i denote the field number and the value of the viscous coefficient at the field i (the field with the number i).

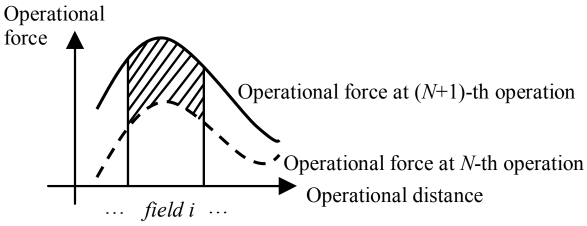

Fig. 2 Profiles (operational force)

If the reward r_i is calculated based on evaluation of operational feeling, $Q(i, d_i)$ reflects the evaluation through learning, here r_i is the reward when the robot passes the field i. In the previous study [5], the evaluation was measured by the degree of convergence between the profiles at the field i obtained in the N-th and (N+1)-th operations. Here, the profile means the plot of the operational force or the velocity of the robot where the operational distance of the robot is assigned to the horizontal axis, as explained in Fig. 2. The degree of convergence at the field i was calculated by the area formed by the plots and the boundary lines of the field i, and a smaller area provided a larger reward at the field i of (N+1)-th operation. This evaluation was based on the assumption that the profile was convergent when a good operational feeling is obtained, and the assumption was based on an experimental observation in the study by Yamada et al. [4]

III. EXTENTION TO EFDA

Operational feeling from the initial position through a goal position is determined by a sequence of viscous coefficients from the first field (the field where the initial position locates) through the goal field (the field where the goal position locates). If we consider extension of the previous study described in the previous section to the case where the number of goal positions which can be selected by an operator is n (n>1), n sequences of viscous coefficients are required and should be learned. In addition, if selection of a goal position is made by an operator in real time, a function of inferring the goal position is necessary, because the goal position cannot be told to the robot in advance before an operation.

To the above problems, the study in this paper considers the extension of the previous study, referred to as EFDA.

1. Sequence of viscous coefficient

When the goal field is the field j, the sequence of viscous coefficients is defined as

$$D_j = \{d_{j1}, \cdots, d_{ji}, \cdots, d_{jj}\}, \tag{2}$$

where $i = 1, \cdots, j$ denotes the field number and $j = 2, \cdots, J$. The field J denotes the field where the farthest goal position from the initial position locates. d_{ji} denotes the viscous coefficient chosen by an agent at the field i when the goal field is the field j.

2. Inference of the goal field

Accurate inference of the goal position is a difficult problem in general. However, adjustment of viscous coefficients is "field-dependent" in the study in this paper. Therefore, the adjustment can be performed if the goal field is inferred. Although several types of algorithms (using Hidden Markov Model, for example) can be considered for inference, this paper considers a simple method.

At the initial stage, an operator is asked to perform *one* operations of positioning to each of all the fields used in a positioning task. By the operations, an initial profile of velocity is obtained for each field, and the robot memorizes the profiles.

In each operation of positioning, the profile of the velocity for the current operation is updated with the measured velocity when the robot enters a new field. The entrance also cues comparison between the updated profile and the profiles memorized in the robot. Integrations of absolute errors between the profiles are calculated, and the field of which profile in the robot gives the smallest value of integration is treated as the goal field at the timing of the entrance. The comparison is repeated until the robot is stopped at the goal field.

After each operation, the profile in the robot is updated. If the robot is stopped at the field k, the profile of the goal field k is replaced with the profile obtained in the current operation. This replacement is important because it is thought that the profile varies according to improvement of skill and fatigue of an operator.

3. Choice of viscous coefficient

If the inferred goal field is the field j when the robot enters the field i, the agent chooses d_{ji} in D_j for control of operational feeling. Furthermore, if the inferred goal field in the next field is the field *j'*, the agent changes choice of the sequence to $D_{j'}$ and select $d_{j'(i+1)}$.

4. Update of the action value function

We assume that the goal field of an operation is the field *k*, and the inferred goal fields from the fields 1 through *k* includes an inference error at the field *i*. If the incorrect inference is the field *k'*, the selection of the viscous coefficients for this operation is $\{d_{k1}, \cdots, d_{k'i}, \cdots, d_{kk}\}$. In this case, the action value function of not $Q(i, d_{ki})$ but $Q(i, d_{k'i})$ is updated for the field *i*.

This update is natural in the context of the theory of reinforcement learning. However, the authors think that an inference error affects learning process. In the above example, if the inferred goal field is correct in the field *i* at the next operation to the goal field *k*, d_{ki} which is not learned in the previous operation to the goal field *k* is selected for the field *i*. This is a problem needed to be discussed. In this paper, the result of experimental investigation is reported in the next section.

IV. EXPERIMENT

1. Setup

Experimental investigation was carried out with the experimental setup as shown in Fig. 1 and some operators. The length of a field was 0.1[m] and seven fields (the fields 1 through 7) were prepared. The task included six goal positions located in the range of 0.2 to 0.7[m] from the initial position at the interval of 0.1[m]. In each of operation, one of goal positions was randomly indicated to the operator by a positioning target. Here, the target position was not told to the robot.

The desired mass and stiffness of the robot were fixed to 10[kg] and 0[N/m], respectively. The desired viscous coefficients prepared for choice by the agent were nine values, ranging from 10 to 50[Ns/m] at the interval of 5[Ns/m]. Adjustment was not applied to the field 1 (including the initial position) and the viscous coefficient of the field (namely, d_{21}, d_{31}, $\cdots$, d_{71}) was fixed to 10[Ns/m], which was the minimum value of the choice. The value was determined based on a well-known result of power-assist devices that a smaller viscous coefficient is preferable in the initial position of an operation.

2. Task

An operator was asked to position the robot at the goal position indicated by the target. After each operation, the operator was also asked to judge whether a preferable operational feeling was obtained or not. If the operator thought that the operational feeling was preferable, adjustment of viscous coefficients to the goal position was finished, and the goal position was excluded from the indication after the next operation.

The operation and the judgment described above were repeated until operational feelings to all the goal positions became preferable for the operator.

3. Results

Table 1 shows the viscous coefficients obtained by one of operators after adjustment. Note again that the field 1 was not included in adjustment. From Table 1, two observations are made. The first is that the values of viscous coefficients around the goal field are larger than those at the field 1. This observation suggests that dissipation of kinetic energy by larger viscosity around the goal field assists positioning of the operator, and the observation is similar to that reported in other studies of power-assist devices. The second observation is that the sequences of viscous coefficients are categorized into two groups: the sequences to the fields 2 and 3, and those to the fields 4 through 7. In the former group, the value of viscous coefficient at the field 2 is 35[Ns/m], whereas its values are 10 and 15[Ns/m] in the latter group.

Table 1 Viscous coefficients after adjustment

goal	field						
field	1	2	3	4	5	6	7
2	10	35	-	-	-	-	-
3	10	35	25	-	-	-	-
4	10	10	20	35	-	-	-
5	10	10	20	25	20	-	-
6	10	15	20	25	20	25	-
7	10	15	15	20	25	25	25

For further consideration of the second observation, correct rates of inference are shown in Table 2. It is observed that high correct rates are marked around the goal field, whereas moderate rates are observed in the field far from the goal field. Especially, the rates at the field 2 when the goal fields are the fields 4 through 7 are in the range of 20.0 to 66.7[%]. This indicates that inaccuracy of inference at the field 2 induces random choice of viscous coefficient, and d_{42}, d_{52}, d_{62} and d_{72} are equally learned to some extent.

Table 2 Correct rates of inference [%]

goal field	field						
	1	2	3	4	5	6	7
2	75.0	100	-	-	-	-	-
3	50.0	66.7	100	-	-	-	-
4	26.7	66.7	80.0	93.3	-	-	-
5	20.0	20.0	80.0	100	100	-	-
6	11.1	22.2	44.4	55.6	77.8	100	-
7	22.2	55.6	77.8	55.6	77.8	100	100

The above consideration suggests that inaccuracy of inference has an influence on learning process. The authors think that the inaccuracy can be taken into account for improved design of the agent, and the design is one of future work.

4. Evaluation of operational feeling

For further study of EFDA in future, the authors attempted to evaluate operational feeling from a viewpoint of energy. The index for evaluation was $J = E_2 / E_1$, where E_1 denotes the energy which an operator exerts on the robot in one operation, and E_2 denotes the energy dissipated by the operator in the operation. The energy is calculated by integration of the product of operational force and velocity throughout the operation. Here, the velocity is always positive because one-way operation is assumed in this paper. Therefore, E_1 and E_2 are obtained by the integrations where the operational force is positive and negative, respectively.

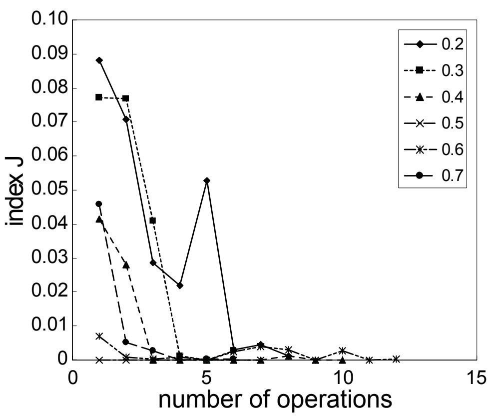

Fig. 3 Transitions of the index J

Figure 3 shows the transitions of J, where the values of $j \times 0.1$ denote the goal position located in the field j. The horizontal axis denotes the number of operations to each goal position. The plots shows that J for all the goal positions become small values (under 5×10^{-3}) at finish of adjustment. However, the plots also indicate that process to finish of adjustment is different between the positions. These observations suggest possibility of measuring preference to operational feeling by energy. However, necessity of another index is also suggested for evaluation of the adjustment process.

V. CONCLUSION

This paper discussed autonomous adjustment of viscosity of a power-assisted positioning task from a viewpoint of operational feeling. Under the assumption that an operator selected one of multi goal positions in real time for an operation, EFDA (Enhanced Field Damping Adjuster) was proposed for realization of preferable operational feelings to all the goal positions, and a function of inferring a goal position and an adjuster using reinforcement learning were developed.

Experimental results showed that adjustment of viscosity was processed so that dissipation of kinetic energy of the robot assisted positioning at the goal position, and also that inaccurate inference of goal positions affected learning process. Considering an improved design of the agent is one of future work.

The study in this paper was supported by Grant-in-Aid for Young Scientists (B) 17760216, and was a part of High-tech Research Center Project (Toyota Technological Institute, 2003-2007).

REFERENCES

[1] R. Ikeura and H. Inoue (1995), Variable Impedance Control of a Robot for Cooperation with a Human, Proc. 1995 IEEE Int. Conf. on Robotics and Automation, 3097-3102

[2] O. M. Al-Jarrah and Y. F. Zheng (1997), Arm-Manipulator Coordination for Load Sharing Using Variable Compliance Control, Proc. 1997 IEEE Int. Conf. on Robotics and Automation, 895-900

[3] T. Tsumugiwa, R. Yokogawa and K. Hara (2002), Variable Impedance Control Based on Estimation of Human Arm Stiffness for Human-Robot Cooperative Calligraphic Task, Proc. 2002 IEEE Int. Conf. on Robotics and Automation, 2002

[4] Y. Yamada, Y. Umetani, H. Daitoh et al. (1999), Construction of a Human/Robot Coexistence System Based on A Model of Human Will - Intention and Desire, Proc. 1999 IEEE Int. Conf. on Robotics and Automation, 2861-2867

[5] T. Morizono, H. Kohno, M. Higashi et al. (2004), A Preliminary Study for Realization of Field Impedance Equalizer with an Automatic Adjusting Function, Proc. 13th IEEE International Workshop on Robot and Human Interaction, 407-412