2025 IEEE/SICE International Symposium on System Integration (SII)
January 21-24, 2025, Munich, Germany

Feedback-Driven Adaptive Task Estimation and Human Error Handling in Human-Robot Collaboration

Kota Tahara, Mihoko Niitsuma, *Member, IEEE*

Abstract— **A human-robot collaboration work system has been developed in which the robot operates following human tasks based on recognition and prediction of human tasks using spatial intelligence. However, the robot's task selection depends on human task recognition, and the robot cannot respond to human error or incorrect recognition of human tasks. In this study, we aim to develop a system that does not depend on the results of human task recognition. The system incorporates human task estimation that accounts for human error and robot motion estimation using reinforcement learning based on human feedback. Its usefulness was evaluated through experiments on human subjects.**

I. INTRODUCTION

In recent years, factories in various industries, including manufacturing, have become increasingly automated. However, industrial robots are not well-suited for complex and precise tasks and lack the flexibility required for high-mix low-volume production. Therefore, Human-Robot Collaboration (HRC), where a person capable of performing complex and precise tasks and flexibly responding to line changes collaborates with a robot suited for long hours of operation and simple repetitive tasks in a shared workspace without safety barriers, is being introduced. In addition, many industrial robots have been developed for the purpose of HRC, called collaborative robots, and their demand is expected to increase in the future.

On the other hand, there are still major issues to be addressed in HRC. The first is safety. Collaborative robots comply with ISO standards [1][2], so safety in case of contact is taken into account, but they are not equipped with safety features such as avoiding conflict with workers. In the current system, workers must constantly be on the alert for the robot's behavior, which is likely to place a heavy workload on collaborative work. In order for workers to focus on their tasks without worrying about the robot's position or task, a safety design that allows the robot to proactively avoid hazards is necessary.

The second issue is work efficiency. Conventional collaborative robot behavior design takes into account the work efficiency of the robot alone, but does not take into account the cooperation with workers, and therefore, smooth and efficient collaborative work has not been realized. Although HRCs are inherently required to work safely and efficiently, there is generally a trade-off between safety and work efficiency. Designing an HRC with an emphasis on one often compromises the other.

Therefore, in recent years, research has been conducted in HRC to change the robot's movement speed and avoid conflicts depending on the distance from humans [3][4], to focus on predicting robot movements [5][6], and to control robot movements [7]. However, issues remain, such as a decrease in the efficiency of the robot behavior itself and a high cognitive load on the worker. In addition, studies in which robots worked in accordance with human tasks recognized from gestures and behavior patterns achieved safe and efficient collaborative work [8][9]. However, one problem is that the robot lacks flexibility in task selection because the prediction results depend on the person's motion trajectory and learning, and the robot performs reactive actions in response to the prediction.

Sanderud et al [10] developed a system that allowed robots to work proactively against hazards by creating a work log based on the time a cell in the workspace was occupied by a worker and predicting the worker's task process. As a result, the worker was able to work without worrying about the robot's position or task. However, because the workspace had to be determined in advance and cells had to be set up, the work area was limited and there was little flexibility in terms of work content.

Therefore, in a previous study [11], we developed a system that enabled a robot to respond flexibly and proactively to a worker by estimating the worker's task process and work time, using the work area and work time automatically extracted from the worker's work. To realize this collaborative work system, spatial memory [12] was applied. Spatial memory is a system that places a Spatial-Knowledge-Tag (SKT), a virtual tag that links a three-dimensional space to electronic data, in the space and enables users to retrieve (access) electronic data by pointing to the tag with a body part. Fig. 1 shows a conceptual diagram of spatial memory. The SKT is automatically placed in each work area extracted from the work data. Work is recognized by retrieving the information stored in each work area according to the worker's actions. However, the robot's task selection depended on the results of human task recognition, and there were issues such as incorrect recognition of human tasks and inability to respond to human error. Therefore, we were able to flexibly respond to various work situations by designing a model to deal with situations that the robot could not actually handle, but this model design required prior knowledge and was skill-dependent.

In this study, we aim to develop a collaborative work system that does not require prior knowledge or rely on the results of human task recognition. We incorporate a human task process estimation function that accounts for human error, along with an automatic generation function for a robot task

* Research partially supported by Chuo University.

K. Tahara and M. Niitsuma are with Chuo University, Bunkyo-ku, Tokyo, 112-8551. (e-mail: niitsuma@mech.chuo-u.ac.jp)

estimation model based on learning from worker feedback. Finally, we demonstrate the usefulness of this system.

II. Collaborative Work System Applying Spatial Memory

In a collaborative work system, this study applies spatial memory to realize the recognition of the worker's work status and the transmission of information necessary for the robot's task selection. The system first generates a model of the robot's behavior using the learning-time system shown in Fig. 2. The robot selects a task from the worker's task, which is recognized using spatial memory and the observed 3D skeletal coordinates of the worker, along with the predicted work task. It then learns the optimal behavior of the robot based on the worker's feedback on the actions performed by the robot and generates a behavior model. The after-learning system shown in Fig. 3 allows the robot to estimate and perform appropriate actions from the generated model. Each block represents each function implemented in the proposed system, which is implemented as a distributed component based on a middleware system called OpenRTM-aist [13].

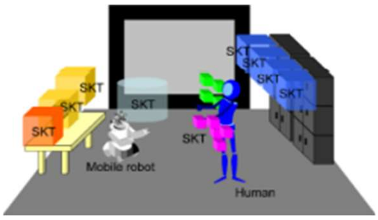

Fig. 1. Conceptual diagram of spatial memory.

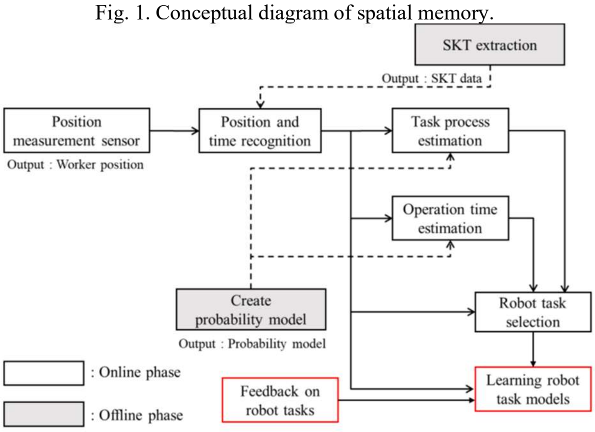

Fig. 2. Overview of the learning-time system, which generates a behavioral model of the robot based on feedback from the worker.

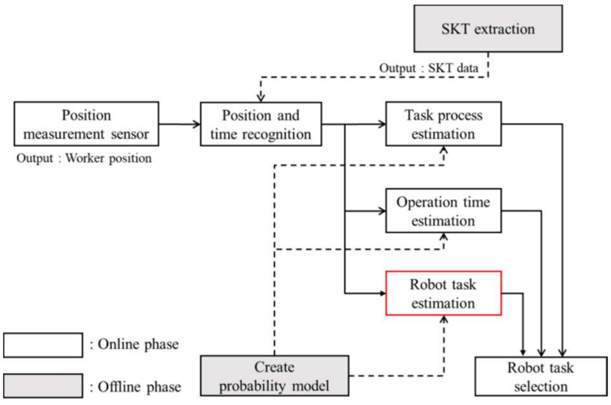

Fig. 3. Overview of the after-learning systems, which allow robots to estimate and perform actions based on behavioral models.

III. Worker's Work Recognition and Work Estimation

A. Extraction of the Worker's Skeletal Coordinates

To recognize the position of the hands and other body parts within the workspace, it is necessary to have a function for estimating the worker's skeleton. This study uses Intel's RealSense and Skeleton Tracking SDK [14] for real-time extraction of worker skeletal coordinates. The SDK is capable of extracting the 3D skeletal coordinates of up to 18 joints every frame in real time. For this study, the coordinates of the worker's right wrist are specifically extracted.

B. Recognition of the Worker's Work Position

Human work in collaborative work is considered to be tied to the position in the workspace. Therefore, clustering is performed on the worker's moving points in the space, and only the potential work areas that are important to the worker are extracted by removing the potential work areas of low importance. In addition, SKTs are generated by clustering the points comprising the extracted work area using a three-dimensional mixed Gaussian model (GMM).

In order for the robot to know where and for how long the worker works, the time when the worker is accessing and not accessing each generated SKT is stored as a work log. In this study, the right wrist of an observed worker is determined to have accessed an SKT when it enters the area of the SKT.

C. Estimation of Task Process and Work Time

In collaborative work with a robot, the next work area is considered to change depending on the task selected by the worker. Therefore, we estimate a worker's task process using a Markov model in which the state transition probability changes depending on the worker's work situation. This state transition model is called a conditional Markov model. The SKT a worker is expected to work on next is estimated by setting a threshold for the work time and adjusting the state transition probability based on the work time. Therefore, human error is recognized by changing the state transition probability so that a state does not transition to the next one when the work time is shorter than a set threshold in the task process estimation.

It also estimates a worker's work time within a given SKT by modeling the transition relationship between each SKT and work time using a hidden Markov model (HMM).

IV. Robot Task Selection Considering Human Error

In this system, when a worker redoes a previous work area due to human error, the robot to erroneously recognizes that additional work is required in that area. In this case, it is expected to transition to other areas in a shorter working time than expected.

The robot's task selection is based on the safety of the worker as the first priority. Tasks to be executed at the SKT where the worker is working and the SKT obtains from the task process estimation results are considered as tasks that may conflict with the worker and are excluded from the executable tasks. Then, based on the number of times each task is executed and the number of times each SKT is accessed by workers, the tasks are prioritized and the task with the highest priority is selected [11]. If the SKT that the worker is accessing and the SKT estimated to work next are the same, the robot waits as a human error. It also waits when there is no executable task.

V. Robot Task Estimation and Task Selection

This study focuses on Markov decision processes (MDPs) to make the robot's task selection independent of the task recognition result. We also estimate the robot's task using the MDP model for motion selection. We consider that the model should be tailored to each worker and be able to be generated by a worker without expertise. Therefore, the robot automatically generates a model by learning the optimal behavior through feedback from the worker on the actions performed by the robot.

A. Robot Task Estimation

The MDP takes a certain state at each time and chooses the actions available in that state. The process then transitions to a new state and receives a reward corresponding to the state transition. The goal of reinforcement learning is to obtain the strategy that maximizes the reward. To optimize a strategy under a given MDP (generated by reinforcement learning as described in section V.B), an objective function is needed, and to compute it efficiently, the objective function must be computed recursively. For this purpose, an equation is needed that relates the objective function in one state s to the next state s'. This equation is called the Bellman equation, and the defining equations are shown in Equations (1) and (2). Two types of objective functions are defined for a measure: the state value function $V^{\pi}(s)$ and the action value function $Q^{\pi}(s,a)$. In this study, the value iteration method is used to obtain the optimal strategy from the Bellman equation and to estimate the behavior.

$$V^{\pi}(s) = \sum_{a} \pi(a|s)\, Q^{\pi}(s,a) \tag{1}$$

$$Q^{\pi}(s,a) = \sum_{s'} P(s'|s,a)\,[R(s,a,s') + \gamma V^{\pi}(s')] \tag{2}$$

where a is the action, π is the policy, P is the state transition probability, R is the reward, and γ is the discount factor.

In this study, we define the SKT currently accessed by the person, the SKT accessed one SKT ago and its work time as state, and the robot's task as action.

B. Generate Task Estimation Model by Q-Learning

In this study, Q-learning is used to generate a Markov decision process model for estimating the robot's task; Q-learning is a reinforcement learning technique that learns by selecting actions in an environment based on the resulting rewards, with the goal of optimizing the state action value (Q-value). The Q-values are updated using equation (3), and learning is performed on the observed states and actions. The learning result of the Q-value is used as the reward in Section V.B. The policy is to use the greedy policy.

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left(R + \gamma \max_{a'} Q(s',a') - Q(s,a)\right) \tag{3}$$

where α represents the learning rate, and the reward R is determined based on the worker's feedback. The worker rewards the observed state and action by providing feedback on the correct robot action for the action performed by the robot based on the task selection algorithm in Chapter IV. However, the worker provides feedback only when the robot performs an incorrect action, because it is considered that the work of the worker to always provide feedback is burdensome and does not allow for smooth collaborative work. In other words, when there is no feedback from the worker, the robot is given positive rewards for its actions. On the other hand, when feedback is given, the robot is negatively rewarded for its actions and positively rewarded for the actions that are given feedbacked. The feedback method is the input of a number (the robot's task number).

C. Robot Task Selection Using Task Estimation

We believe that the robot's task estimation results are applicable to a variety of situations, independent of the accuracy of access to SKTs. Therefore, the task to be performed by the robot is determined by combining the estimation results of SKTs that require robot task support obtained in Section V.A., the number of robot behaviors in each SKT, and the estimation results of task process and work time.

VI. Experiment

Through comparative experiments with a robot that performs task selection based on observed worker tasks and estimated tasks (the before-learning system), we verified whether the robot could perform optimal actions based on human teaching and the validity of the feedback method. This experiment was conducted based on the approval of the Ethical Review Committee for Research on Human Subjects at Chuo University.

A. Experimental Conditions

In this experiment, an experimental task was designed to simulate an assembly operation. The actual experimental environment is shown in Fig. 4. The sensor that observed the worker's right wrist was mounted on the ceiling and observed from directly above the worker. The overall goal of one task was for the worker to create four assemblies consisting of 7 and 14 blocks shown in Fig. 5, and for the robot to carry the assemblies to Storage. The order in which the four assemblies were created was not specified; they were assembled in the order the worker chooses to create them. The worker assembled the blocks collected in WS1 in WS2 and placed the

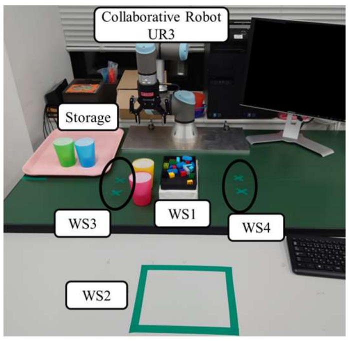

Fig. 4. Experiment environment.

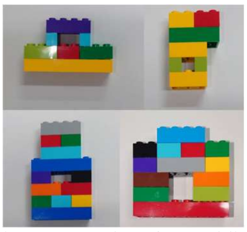

Fig. 5. Examples of assemblies with 7 blocks (top) and 13 blocks (bottom).

finished product in a container. Then, the worker carried 7 assemblies to WS3 and 13 assemblies to WS4 (see Fig. 4). Since the work time is expected to vary with the number of blocks, the state transition probability of the conditional Markov model is expected to change at this time. To achieve this goal, the robot was expected to perform the tasks of refilling blocks and empty containers at WS1 and retrieving finished products at WS3 and WS4 twice each. Therefore, the task achievement rate is defined as the ratio of the total number of assemblies actually created by the worker and the total number of finished assemblies carried by the robot to Storage to the total number of assemblies to be created and carried. The subjects were 10 students (9 males and 1 female) with an average age of 23 ± 1.04 years. Each subject worked for approximately 4 hours in total, with each experimental task lasting around 5 minutes.

In this experiment, a training experiment was also conducted to generate a task estimation model for the robot. When a robot using the before-learning system performed an incorrect action and the worker noticed it, the worker stopped the task and entered the correct action number of the robot. In this case, if the teaching robot is not taught at the appropriate time, incorrect learning may occur, but the subject may not notice the robot's incorrect action. Only when the worker noticed this during the process, the subject reported it to the engineer, who corrected the learning content. The parameters used were a learning rate of 0.1 and a discount rate of 0.9.

B. *Experimental Procedure*

The following procedure was used to conduct the experiment. (1) The worker performed the experimental task eight times under remote control by a skilled operator. (2) Based on the data obtained in (1), generated SKTs and determined the threshold for task process estimation. (3) Performed the experimental task with the before-learning system 8 times. (4) Performed the training experiment task eight times to generate a task estimation model for the robot. (5) Performed the experimental task with the after-learning system eight times. (6) The experiment was completed by answering a subjective evaluation questionnaire with free-response questions. After each experimental task, participants responded to a cognitive load evaluation questionnaire called NASA-TLX [15] and an impression evaluation questionnaire after the experimental tasks with the pre-learning system and the after-learning system. The order in which the experimental conditions were presented was fixed to ensure that the subjects could become sufficiently familiar with the task during the teleoperated experimental task and understand, through the task in the before-learning system, that the robot performed incorrect behavior. In addition, sufficient breaks were provided between each experimental task.

VII. Results and Discussion

A. *Results for All Subjects*

The results of access accuracy to SKTs and task achievement rates for the before- and after-learning systems are shown in Fig. 6 and 7. Access accuracy is defined as the ratio of the number of correct access decisions to the number of times a worker actually worked on each SKT. The results showed that there was no significant difference between the before- and after-learning systems in both access accuracy and task achievement rate. The results suggest that the intended learning effect of human teaching was not achieved.

Table I shows the number of times a worker should give feedback, the number of times the worker actually gave feedback, and the number of times the worker reported to the technician and corrected the learning content during the eight training tasks. The results indicated that workers provided feedback a large percentage of the time relative to the number of times they should have provided feedback. Thus, the task achievement rate was expected to improve significantly. Therefore, since we confirm that the task achievement rate is sufficiently high in the before-learning system for different workers and that there are differences in the number of times feedback is given, the following discussion is divided into two groups: one with high task achievement rates and the other with low task achievement rates in the before-learning system. Groupings were made by non-hierarchical clustering (k-means clustering) on the results of task achievement rates in the before-learning system. The results showed that six students were classified in the group with the higher task achievement in the before-learning system and four in the group with the lower achievement.

B. *Access Accuracy and Task Achievement Rate*

The access accuracy and task achievement rate results for the group with the higher achievement are shown in Fig. 8, and the results for the group with the lower achievement are shown in Fig. 9. For the higher-achieving group, there were no significant differences in either access accuracy or task achievement rate. On the other hand, the task achievement rate was significantly higher in the after-learning system than in the before-learning system, even though the access accuracy

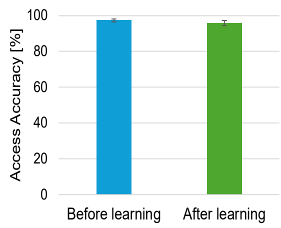

Fig. 6. Access accuracy.

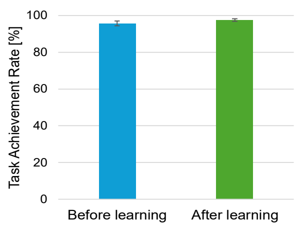

Fig. 7. Task achievement rate.

TABLE I. Results of Training Experiments for All Subjects

	Average	Standard deviation
Number of times feedback should be given	5.6	4.84
Number of feedbacks	4.6	4.84
Number of revisions	0.5	0.92

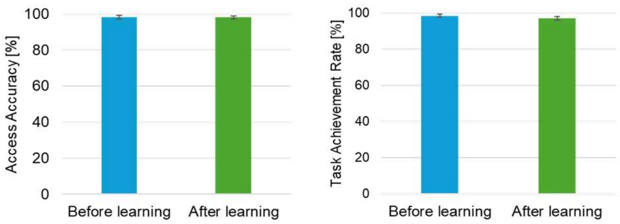

Fig. 8. Results for the groups with the higher task achievement.

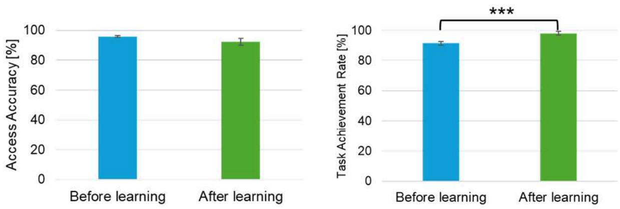

Fig. 9. Results for the groups with the lower task achievement (***: $p < 0.001$).

was lower in the lower-achieving group (no significant difference). This suggested that for those with lower task achievement, robot teaching had a learning effect, and that the system was independent of access accuracy.

There were three main reasons why the robot failed to achieve its goals using this system: first, the robot was unable to correctly recognize human tasks, which was confirmed by eight subjects. Specifically, the size of the generated SKT was too small to be accessed, and the reproducibility of the task was low and the way the task was performed was different from the time when the SKT was generated. However, it was confirmed that this situation could be learned by the worker's teaching. The second was false access to SKTs, which was confirmed in three subjects. Specifically, it was confirmed that the subjects unintentionally accessed other SKTs depending on the proximity between SKTs and the way they worked. This situation was not manageable through teaching, as the system was designed to respond when teaching-based learning failed to identify the task. However, we believe that this problem may be solved by improving the placement of SKTs and the access judgment method. Third, human error occurred after working for more than the threshold time to recognize human error, which was confirmed by two subjects. This situation is not capable of being handled by teaching robots, but it may be improved if robots are capable of correctly detect human errors. This indicated that this system could supplement the situation in which a worker was unable to recognize a task by human instruction. Since the main cause of a robot's inability to achieve its goal was its inability to recognize the work situation, we considered the teaching robot learning method to be appropriate for this situation. However, the problem was that the after-learning system was not capable of handling unteaching situations and was unable to re-teach at that time.

C. *Training Experiments*

Table II shows the results for the higher-achieving group and Table III shows the results for the lower-achieving group in the training experimental task. The results indicated that there was no significant difference in the task achievement rate among the higher-achieving group because they received less feedback. On the other hand, the lower-achieving group received more feedback, which allowed them to learn

TABLE II. RESULTS OF TRAINING EXPERIMENTS FOR THE GROUPS WITH THE HIGHER TASK ACHIEVEMENT

	Average	Standard deviation
Number of times feedback should be given	2.8	2.1
Number of feedbacks	2.5	2.3

TABLE III. RESULTS OF TRAINING EXPERIMENTS FOR THE GROUPS WITH THE LOWER TASK ACHIEVEMENT

	Average	Standard deviation
Number of times feedback should be given	9.6	4.8
Number of feedbacks	8	5.5

appropriate behavior in situations where they could not recognize the task and significantly improved their task achievement rate. It could also be seen that workers were able to actually provide feedback a large percentage of the time in relation to the number of times they should have provided feedback. Therefore, the feedback method using numerical input was considered to be appropriate.

C. *Questionnaire*

The results of the impression evaluation questionnaire are shown in Fig. 10 and 11. For the higher-achieving group, there was no significant difference between the before- and after-learning systems in any of the items. On the other hand, for the group with lower grades, the after-learning system gave significantly better impressions on all items.

The results of the impression evaluation questionnaire are shown in Fig. 12 and 13. For the higher-achieving group, there was no significant difference between the before- and after-learning systems in any of the items. On the other hand, for the group with lower grades, the after-learning system gave significantly better impressions on all items.

The results of the subjective evaluation questionnaires, including free-responses, are shown in Fig. 14 and 15. It was considered that the higher-achieving group felt the feedback

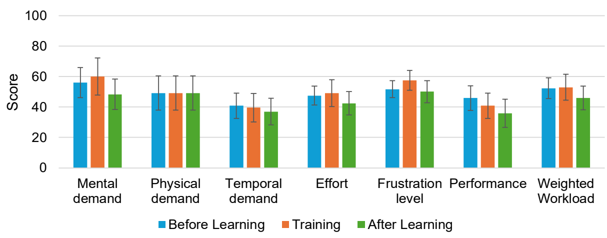

Fig. 10. NASA-TLX results for the groups with the higher task achievement.

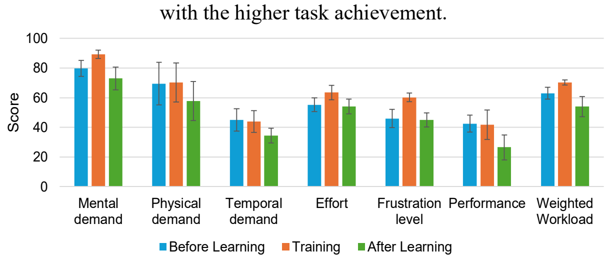

Fig. 11. NASA-TLX results the groups with the lower task achievement.

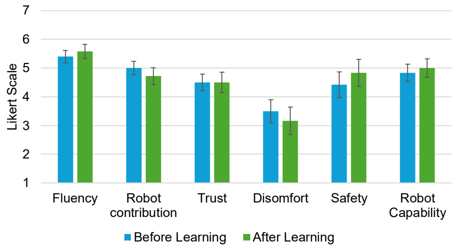

Fig. 12. Results of the impression evaluation questionnaire for the groups with the higher task achievement.

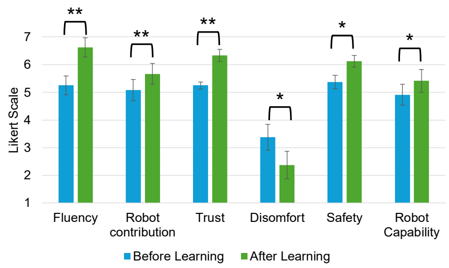

Fig. 13. Results of the impression evaluation questionnaire for the groups with the lower task achievement (*: $p < 0.05$, **: $p < 0.01$).

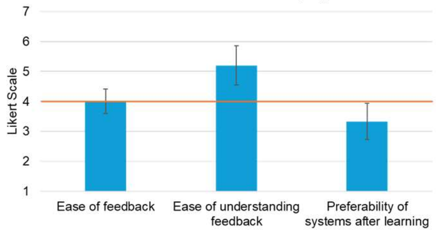

Fig. 14. Results of the subjective evaluation questionnaire for the groups with the higher task achievement.

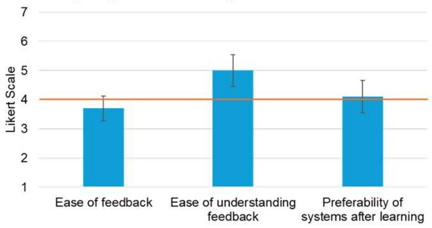

Fig. 15. Results of the subjective evaluation questionnaire for the groups with the lower task achievement.

task was easier than the lower-achieving group, because the number of feedbacks was smaller. However, both groups felt that the work was difficult because they were conscious of the robot's movements, and did not receive a high evaluation. Regarding the understandability of the feedback task, both groups of subjects rated it highly because they only had to enter numbers, and there was no significant difference between the groups, indicating that they found the feedback task easy to understand. On the other hand, there were some subjects who were so focused on their work that they found it difficult to notice any errors in the robot's behavior. Therefore, it was considered necessary to have a function that allowed the robot itself to notice and respond to errors and a function that told the worker when to teach. In addition, when asked whether they preferred the after-learning system to the before-learning system, the group with the lower grades preferred the after-learning system because they felt the learning was more effective. However, the higher-achieving group preferred the pre-learning system because they did not feel the need to be taught because the learning effect was small.

VIII. Conclusion

In this study, we added a task estimation function for workers that takes human error into account and a task estimation function for the robot that uses learning by teaching robot by workers. As a result, the effect of learning by teaching was obtained for workers for whom the before-learning system was not sufficient. The system was then able to supplement the worker's inability to recognize the task, which made the robot's task selection independent of access accuracy.

In the future, we aim to construct a system that enables constant teaching, and to add functions that allow the robot to recognize and respond to errors autonomously while communicating the timing of teaching to the worker.

Acknowledgment

This research was supported by the Chuo University Joint Research Grant.

References

[1] ISO 10218-1:2011, Robots and robotic devices - Safety requirements for industrial robots - Part 1: Robots.

[2] ISO 10218-2:2011, Robots and robotic devices – Safety requirements for industrial robots - Part 2: Robot systems and integration.

[3] L. M. Amaya-Mejia, et al., "Vision-Based Safety System for Barrierless Human-Robot Collaboration," 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Kyoto, Japan, pp. 7331-7336, 2022.

[4] J. Park, et al., "A Digital Twin-based Workspace Monitoring System for Safe Human-Robot Collaboration," 2022 10th International Conference on Control, Mechatronics and Automation (ICCMA), Belval, Luxembourg, pp. 24-30, 2022.

[5] X. Zhao, et al., "An Actor-Critic Approach for Legible Robot Motion Planner," 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, pp. 5949-5955, 2020.

[6] S. Sonawani, et al., "Projecting Robot Intentions Through Visual Cues: Static vs. Dynamic Signaling," 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Detroit, MI, USA, pp. 7931-7938, 2023.

[7] Y. Ye, H. You and J. Du, "Improved Trust in Human-Robot Collaboration With ChatGPT," in IEEE Access, vol. 11, pp. 55748-55754, 2023.

[8] J. Male, G. A. Al, A. Shabani and U. Martinez-Hernandez, "Multimodal sensor-based human-robot collaboration in assembly tasks," 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 1266-1271, 2022.

[9] A. M. Zanchettin, A. Casalino, L. Piroddi and P. Rocco, "Prediction of Human Activity Patterns for Human-Robot Collaborative Assembly Tasks," in IEEE Transactions on Industrial Informatics, vol. 15, no. 7, pp. 3934-3942, July 2019.

[10] Audun Rønning Sanderud, M. Niitsuma, "Responsible Robots - A Novel Approach to Safe and Productive Human-Robot Collaboration, " Chuo University, 2016, Ph. D. thesis.

[11] K. Tahara, et al., "Human-Robot Collaboration System Based on Task Estimation Using Spatial Memory in Intelligent Space and Robot Task Selection by Markov Decision Process," ROBOMECH2023 in Nagoya, 2P1-D05, 2023 (in Japanese).

[12] M. Niitsuma, et al., "Spatial Memory as an Aid System for Human Activity in Intelligent Space," IEEE Trans. on Industrial Electronics, vol. 54, no. 2, pp.1122-1131, 2007.

[13] "OpenRTM-aist," https://openrtm.org/openrtm/, (accessed Jan. 26, 2023).

[14] Intel, "Skeleton Tracking SDK by cubemos ™ - Intel RealSense Technologies," https://www.intelrealsense.com/skeleton-tracking/, (accessed Jan. 29, 2023).

[15] S. Haga and N. Mizukami, "Japanese version of NASA Task Load Index: Sensitivity of its workload score to difficulty of three different laboratory tasks," The Japanese Journal of Ergonomics, vol. 32, no. 3, pp.71-79, 1996.