Automatic action recognition algorithm for industrial manual workers with human skeleton and object information

Time and motion studies, which involve dividing a worker's actions into several micro-actions, are a fundamental analytical technique in industrial engineering (IE). Through this analysis, we can identify redundant parts of the worker's actions and make improvements based on it [1]. However, the problem with this approach is that it takes a large amount of time since it is usually performed by human analysts by hand. Therefore, to address this problem, we developed an automatic action recognition algorithm for industrial manual workers.

Although there are a huge number of studies that have tried to develop human action recognition algorithms [2], there are a relatively small number of studies that have dealt with human action in industrial situations. This is because, in the context of industrial manual workers, object and human-object interaction information must also be considered to recognize the worker's action. For example, although the action of reaching for an object and transporting an object are similar in terms of human kinematics, we can clearly discriminate between them in terms of the interaction of the objects. Therefore, to address these issues, we developed a specific algorithm which can deal with the human object interaction information.

Figure 1 shows the overview of the proposed algorithm's architecture. The input of the network is the image data of the human motion in the target task. This data is sent to both the skeleton recognition algorithm and object detection algorithm, and we get both human skeleton and object position (bounding box) information. Then, each data is input to the LSTM network, a type of machine learning technique for processing time series data, and the algorithm finally obtains an action label for each time sequence. Hence, by processing human motion information and object information, we can consider the interaction between the human and the objects.

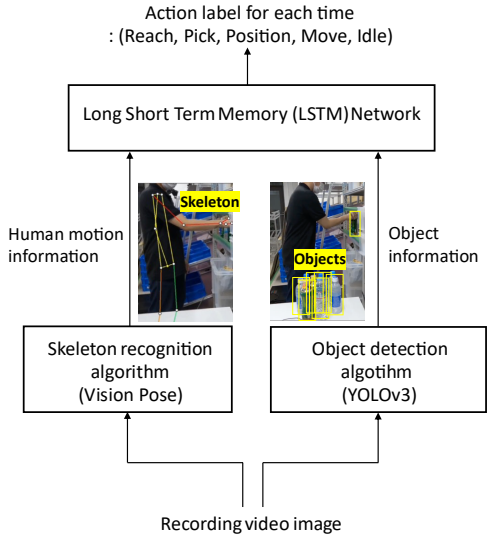

Figure 1. Overview of the proposed algorithm.

To verify the effectiveness of the proposed algorithm, we conducted an experiment. In the experiment, a participant performed a picking and placing task by moving target objects from a shelf to a desk, or vice versa. The motions of the participant were recorded by a RGB video camera. Figure 2 shows the results of the experiment. As shown in Figure 2, we can describe the participant's action well at each time sequence and it bring us some insights to improve the process. From Figure 2, we can identify motions that can be performed relatively quickly (good action) and those that take relatively longer (redundant action). The mean accuracy of recognition for this experimental task is around 90%. Since we only verified the performance of the proposed recognition algorithm in an experimental environment, we plan to verify it in more practical situations in the future.

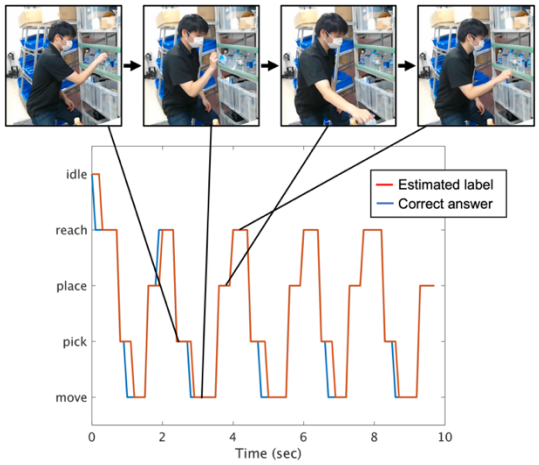

Figure 2. Results of the experiment.

Keywords: Action Recognition, Object Detection, Industrial Engineering, Time and Motion Study.

References

[1] Barnes, R. M. (1991). Motion and time study: design and measurement of work. John Wiley & Sons.

[2] Takamido, R., & Ota, J., (2022). "Robot system design based on the unreliable human demonstration data". Proceedings in the 40th annual conference of the Robotics Society of Japan (RSJ), RSJ2022AC3F1-05, (pp. 1-4). (in Japanese)

[3] Zhang, H. B., Zhang, Y. X., Zhong, B., Lei, Q., Yang, L., Du, J. X., & Chen, D. S. (2019). A comprehensive survey of vision-based human action recognition methods. *Sensors*, 19(5), 1005.