Development of the Control Method of the Walking Assist Robot Using Reinforcement Learning

Shuhei AYUKAWA[a], Hajime TAKADA[b] and Yoshifusa MATSUURA

[a] *Division of Systems Integration, Graduate School of Engineering, Yokohama National University, Japan*
[b] *Division of Systems Research, Faculty of Engineering, Yokohama National University, Japan*

Abstract. In this study, the development of the assist robot to help the walk and to avoid falling in advance of a person whose foot is weak as the elderly is targeted. First of all our study, the model of the assist robot and user were made in the simulator using the physics engine called Open Dynamics Engine (ODE) [1]. By repeating the learning of millions of times, this model will continue to seek the best action. In the definition of "action" and "state", how to provide the case analysis is a difficult point. It is necessary to consider the case divided while performing trial and error by examining and measuring the walking data of actual human and value of the torque exerted at leg joint. In the present circumstances, it was measured load distribution of the sole part of the robot by three sensors. We have to validate the method that the status is defined from that load distribution, and the joint of the robot is applied torque according to the state. Finally it is the goal to create the assist robot actually, and we evaluate the performance of control was developed.

Keywords: Walking, Assist, Robot, Reinforcement Learning, Welfare Technology
PACS: 87. Biological and medical physics

BACKGROUND

Currently, as a walking assist robot of available power supply, "HAL" of CYBERDYNE is famous. It has very high performance, but there is a big drawback. Due to perform an assist by analyzing bio potential signal of humans in HAL, it takes a long time to the mounting, and it is very expensive. In response to it, in our study, we develop the less expensive and adapting to individual differences assist robot by creating a control algorithm of the assist by using a reinforcement learning.

SIMULATION

This time, in order to create the control algorithm, Q-learning algorithm is used, which is one of the reinforcement learning. In Q-learning, the robot takes action in accordance with some measures. It receives a reward from the environment for its actions, and updates the action-value function (Q value) defined by the state and action. Updating expressions Q value is expressed by the following.

$$Q(s_t, a_t) \leftarrow (1-\alpha)\cdot Q(s_t, a_t) + \alpha \cdot \Delta Q \quad (1)$$

$$\Delta Q = r_{t+1} + \gamma \max_a Q(s_{t+1}, a_t) \quad (2)$$

Q : Q value *s : state* *a : action* *r : reward* *α : learning rate* *γ: discount rate*

Optimal behavior is determined by combination of Q values that was finally acquired. Q-learning is characterized by not requiring prior information such as the surrounding environment or detailed data analysis as was needed by the assist algorithms developed to date are.

First we went to build a model with reference to the figure of the elderly (Male over the age of 60) [2] to be the target in ODE. The model is divided into head, body, upper arm, forearm, thigh, lower leg, the foot and equivalent weight [3] is set for each. The various parts are joined by a hinge joint, and it can perform speed control. Assist robot is assumed to be 12kg on the total weight of both feet. In the model, it is possible to measure such as the centroid position, the inclination of the body and the values of the load of three sensors installed in the sole.

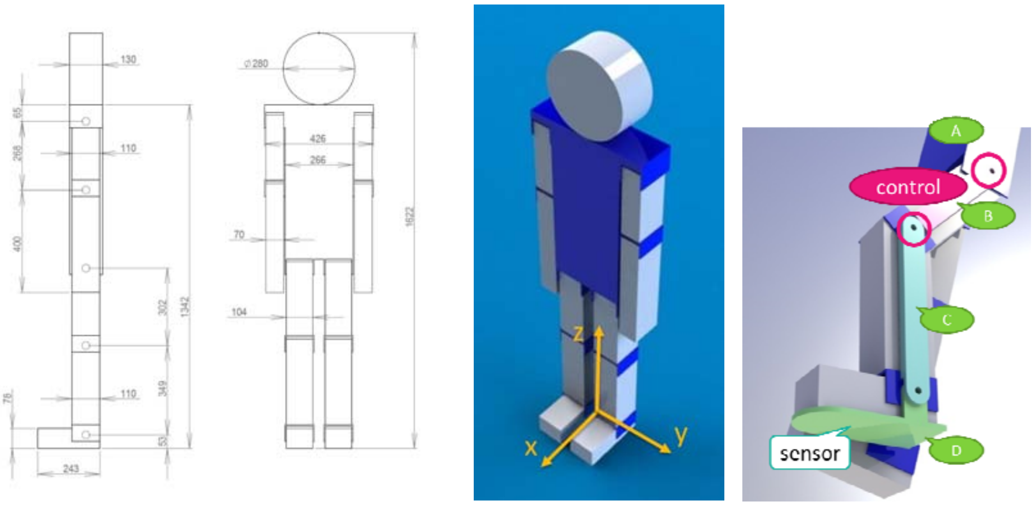

FIGURE 1. Size and schematic of the simulation model

TABLE 1. Equivalent weight of the elderly and the assist robot

Parts(elderly)	*Weight [kg]*	*Parts (assist robot)*	*Weight [kg]*
head	4.6	waist (A)	3.0
body	28.8	thigh (B)	1.5
upper arm	1.6	lower leg (C)	1.5
forearm	1.4	sole (D)	1.5
thigh	6.0		
lower leg	3.2		
foot	1.1		

In the current simulation, as a definition of the state, it is learning and evaluated by using a combination of two indices of the ground state of the sole and the inclination degree of the trunk. In the future, we are considering the state division method using the value of the three load sensors of the sole. As the action, by splitting the amount of torque the assist robot is applied to the legs as shown in Table 2, we are doing an action defined by that combination. In this case, amount of torque applied is less than 25% compared from exerting joint torque in the elderly [4], it is limited to a sufficiently small value. At the end, we provide penalty when the model fall as the reward. In addition to this, we are rewarded in accordance with the height of the head position, smallness of the trunk tilt, the length of the duration of the non-fallen state and the stability of the barycentric velocity. We compare the results due to differences in these rewards and exploring how to give the best reward.

TABLE 2. Splitting method of the action

	Additional torque (y-axis of FIGURE 1)										
hip joint	18	15	12	9	6	3	0	-3	-6	-9	-12
knee joint					6	3	0	-3	-6	-9	

PRODUCTION AND EXPERIMENT

In addition to creating the control algorithm by simulation, we have performed in parallel production of assist robot to adapt the control. At the waist portion of the assist robot, we have tied a body made of a battery. The link extending along the leg from the waist is driven by servo motors attached to the knee joints and hip. The motors are reduced by the planetary gears. Further we have installed three load sensors per one leg on the foot bottom. We measure the load there and use to control. Load sensors are used to load cell (LMB-A (500N/1kN)) of Kyowa Electronic Instruments.

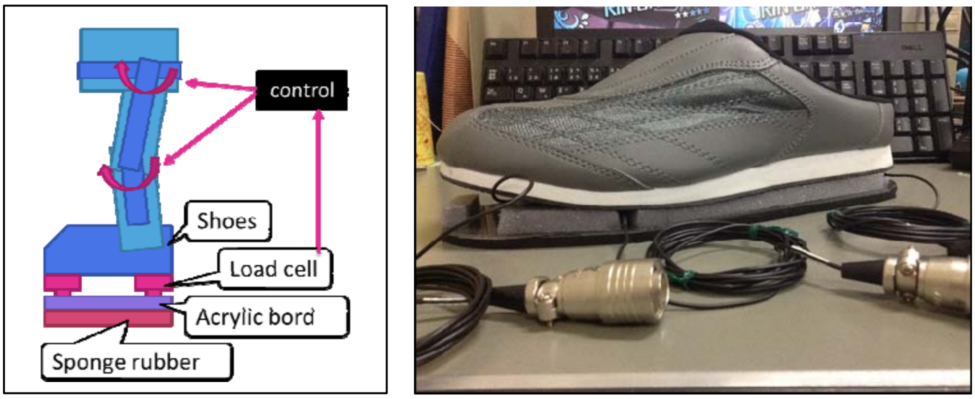

FIGURE 2. Schematic of assist robot and the foot bottom

The up to now, we have production of the sensor section to put on the back foot. The mounting position of the three sensors is set on the base of the thumb, the base of the little finger, the position of the heel. They are located where the load is applied particularly large when human walks. In actual equipment that was produced, it has been confirmed that each sensor is outputting an accurate measurement.

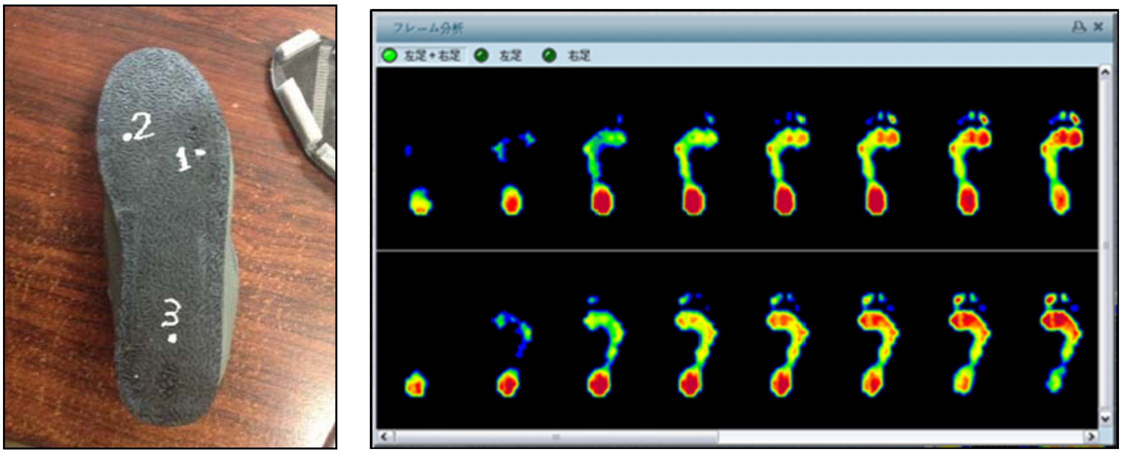

FIGURE 3. Sensor position and load distribution at the time of the walk [5]

Hereafter, in parallel with the creation of the control algorithm by simulation, we will do fall detection experiments with the sole sensor. Here, we will reproduce the fall by tripping which is common in simulator and falling of elderly in the state that I put a foot sensor, and measure the behavior of the sensor value at that time. By extracting the difference between just before falls and the stable walking therefrom, it is the purpose to obtain an indication of the state division in simulation.

REFERENCES

1. Robot Simulation, Demira Kiminari, 2007
2. Research on the Elucidation of the Characteristics of the Elderly for the Design of Elderly Equipment, Elderly Development Project
3. Health Basics Section, Bodywise, 2003
4. Human Characteristics Database, Institute of Technology and Evaluation, 2013
5. GRAND SHAKER, UMEMIYA