A Human-Machine Interaction System:

A Voice Command Learning System Using PL-G-SOM

Takashi KUREMOTO
Graduate School of Science and Engineering
Yamaguchi University
Ube, Japan
wu@yamaguchi-u.ac.jp

Tetsuya YAMANE
Dept. of Information Science and Engineering
Yamaguchi University
Ube, Japan
doramaga@hotmail.com

Liangbing FENG
Graduate School of Science and Engineering
Yamaguchi University
Ube, Japan
n007we@yamaguchi-u.ac.jp

Kunikazu KOBAYASHI
Graduate School of Science and Engineering
Yamaguchi University
Ube, Japan
koba@yamaguchi-u.ac.jp

Masanao OBAYASHI
Graduate School of Science and Engineering
Yamaguchi University
Ube, Japan
m.obayas@yamaguchi-u.ac.jp

***Abstract*—This paper proposes a voice command learning system for partner robots acquiring communication ability with instructors. Parameter-less Growing Self-Organizing Map (PL-G-SOM), an intelligent pattern recognition model given by our previous work, is used and computational feeling of robots is also adopted to improve the human-machine interaction system. AIBO robot was used in the experiment and the results of real environment showed the effectiveness of the proposed methods.**

Keyword—partner robot; PL-G-SOM; human-machine interaction; computational feeling

I. INTRODUCTION

Partner robots, which are robots work in the ordinary environment of human society such as pet robots, entertainment robots, service robots and so on, appear in recent years vigorously. When we design and develop this kind of robots, an important function needs to be considered is the ability of communication with their owners. Recently, we proposed a voice command recognition/learning system named Parameter-Less Growing Self-Organizing Map (PL-G-SOM) [1] which improved conventional Kohonen's SOM [2] with variable topology and less number of parameters. The effectiveness of PL-G-SOM was confirmed by a voice instruction learning system [1] composed by multiple layers [3], however, stayed on the stage of offline execution. In this paper, we apply PL-G-SOM to a pet robot "AIBO", a product of Sony Ltd. 2003 [4], to build an online voice command recognition and action learning system. The development of the real robot command learning system used OPEN-R SDK with C++ language, i.e., an application software framework provided by AIBO's website [4]. The experiment results showed that the average success rate of voice command recognition reached the level of what of PL-G-SOM in simulation experiments reported by [1].

II. PL-G-SOM AND A VOICE INSTRUCTION LEARNING SYSTSEM OF A ROBOT

A. *Parameter-Less Growing Self-Organizing Map (PL-G-SOM) [1]*

Parameter-Less Growing Self-Organizing Map (PL-G-SOM) [1] was proposed by our previous work. The theoretical description is given as follows.

Generally, Kohonen's SOM algorithm [2] maps *n*-dimension feature data in an input space $\mathbf{x}(x_1, x_2\ ...,\ x_n)$ to a unit i in a low-dimensional output space with connections $\mathbf{m}_i(m_1, m_2\ ...,\ m_n)$ by a simple rule using Euclidean distance, winner-takes-all:

$$\boldsymbol{c} = \arg\min_i(\|\mathbf{x} - \mathbf{m}_i\|), \qquad (1)$$

i.e., a high dimensional input is corresponded to a most suitable unit i with position $\boldsymbol{c}$, best-match-unit (*BMU*) on the output map, $i=1, 2, ..., k \leq N\times M$. For all inputs and initial connections with random values, a competitive learning rule enhances that the input data with similar features keep closely on the visualized topological output map:

$$\begin{aligned}\mathbf{m}_i(t+1) &= \mathbf{m}_i(t) + \Delta\mathbf{m}_i(t) \\ &= \mathbf{m}_i(t) + \varepsilon(t)h_{ci}^{\varepsilon}(t)(\mathbf{x}(t) - \mathbf{m}_i(t))\end{aligned}. \qquad (2)$$

Here, $\varepsilon(t)$ is an adaptive learning rate, and $h_{ci}^{\varepsilon}(t)$ is a neighborhood function, $\sigma(t)$ is the neighborhood size. All of them are calculated by the distance between input and the *BMU*:

$$\varepsilon(t)=\frac{\|\mathbf{x}(t)-\mathbf{m}_c(t)\|^2}{r(t)}, \tag{3}$$

$$r(t)=\mathbf{max}(\|\mathbf{x}(t)-\mathbf{m}_c(t)\|^2, r(t-1)), \tag{4}$$

$$r(0)=\|\mathbf{x}(0)-\mathbf{m}_c(0)\|^2, \tag{5}$$

$$h_{ci}^{\varepsilon}(t)=\exp\left(-\frac{\|\mathbf{c}_i-\mathbf{c}\|^2}{(\sigma(\varepsilon(t)))^2}\right), \tag{6}$$

$$\sigma(\varepsilon(t))=\sigma_{\max}\cdot\varepsilon(t),\ \ \sigma(\varepsilon(t))\geq\sigma_{\min}. \tag{7}$$

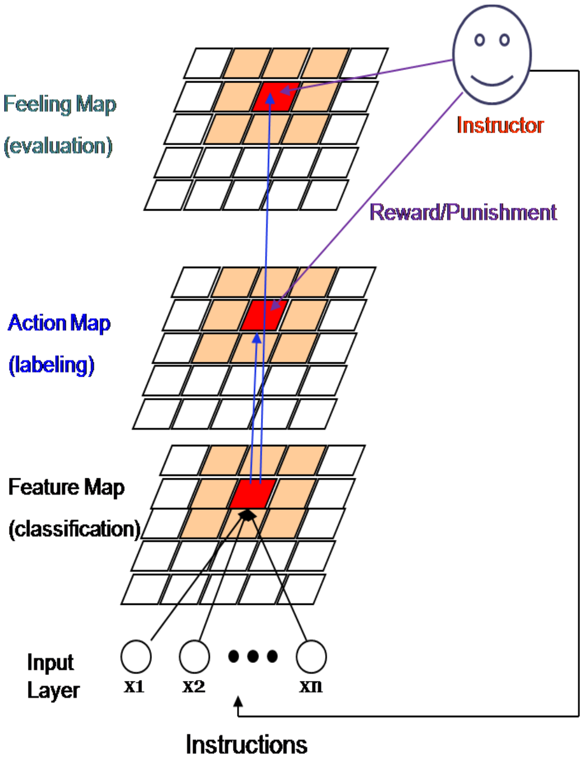


Figure 1. A voice instruction learning system using PL-G-SOM and reinforcement learning algorithm given by Kuremoto et. al [1].

where $\sigma_{\max}, \sigma_{\min}$ are positive parameters, for example, the value may be the size of the map and 1.0, respectively. c_i, c denotes the positions of an arbitrary unit on the output map and *BMU*, respectively, *i=1, 2, …, k* ≤ *N*×*M*. Obviously, $h_{ci}(x)\geq 0$, $h_{ci}(0)=1$, $h_{ci}(\infty)=0$.

To overcome the limitation of the fix size of map, growing SOMs (G-SOMs) are proposed by [8]-[11]. A small size of the feature map is set initially, and when a new input is not able to find a *BMU* from the initial map, a new row/column is inserted in to enlarge the feature map. For example, a new node *r* in the new row/column is inserted into the middle of node *c* and node *f*, where *c* is the nearest node to the new input and *f* is the neighbor of *c*. The weight of connection between the input and the new node has an average value of *c* and *f*,

$$\mathbf{m}_r=0.5(\mathbf{m}_c+\mathbf{m}_f), \tag{8}$$

for nodes which are *r*'s neighbors in the new row/column:

$$\mathbf{m}_{r+l}=0.5(\mathbf{m}_{c+l}+\mathbf{m}_{f+l}). \tag{9}$$

Where *l=1, 2, …, N* or *M*. Unit *f* is chosen which has a largest Euclidean distance from the *BMU*-like *c* among the neighbors of *c*, and after this process, the map size changes to *N*×(*M*+1), or (*N*+1)×*M*.

B. A Voice Instruction Learning System

The PL-G-SOM is adopted into a voice instruction learning system, as shown in Figure 1 in [1], for an intelligent robot.

Feature Map is a PL-G-SOM, each unit of Action Map and Feeling Map corresponding to the growing Feature Map.

The value of units on Action Map is given by a value function of state and action, i.e. (10), where $Q(s_t, a_t(i))$ has the value of selected action $a_t(i)$ of the robot, at the state s_t which is an input voice command. Here, $Q(s_0, a_0(i))$ is set by random numbers initially.

TABLE I. MAIN SPECIFICATIONS OF THE ROBOT AIBO

CPU	64bit RISC 576MHz
RAM	64MB SDRAM
Program media	Dedicated the AIBO robot "Memory Stick™"
Degree of freedom	20 (Head 3; Mouth 1; Legs 3x4; Ears 1x2; Tail 2)
Image input	350,000-pixel CMOS image sensor
Audio input	Stereo microphones
Audio output	Speaker
Sensors	Infrared distance sensors 2; Acceleration sensor; Touch sensors (Head, Back, Chin, Paw)
Wireless LAN	IEEE802.11b
Power consump.	About 7W (in standard mode)
Operating time	About 1.5 hours
Power charging	About 2.5 hours
Action Condition	temperature / humidity : 5- 35℃ / 10-80%
Dimensions	About180x278x319 mm (W x H x D)
Weight	About1.6kg (including battery pack)

$$Q_{t+1}(s_{t+1}, a_{t+1}(i))=Q_t(s_t, a_t(i))\pm r. \tag{10}$$

Where $\pm r$ is the empirical value of reward (+)/punishment (-) given by the instructor, for example, a positive constant when the robot acted correctly according to its policy function (11) and a negative constant oppositely.

$$\pi_t(a_t(i)\,|\,s_t)=\frac{e^{\frac{Q_t(s_t,a_t(i))}{T}}}{\sum_{j\in A} e^{\frac{Q_t(s_t,a_t(j))}{T}}}. \tag{11}$$

π is a probability of action excitation. Here T is a positive parameter named temperature [5], higher T causes an active

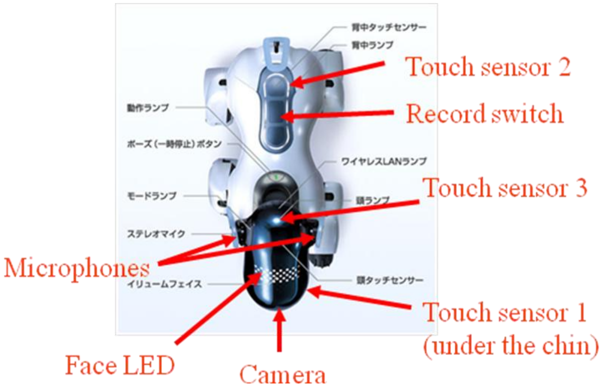


Figure 2. A robot used in the Experiment: AIBO (Sony Ltd., ERS-7, 2003).

exploration of actions (each action is selected more randomly) and lower T gives a greedy selection of the action with higher Q value.

Partner robots, either pet robots or other kinds of entertainment robots, are expected to express what they are "thinking" during interaction with people. To represent "feeing" of a robot, we proposed a feeling formation model in [6]. In this study, the computational feeling model is used to renew Feeling Map of voice command learning system (See Figure 1). The distance from input pattern to *BMU* of Feature Map and the reward from instructor are used to calculate feeling values which is normalized in [-1.0, 1.0] where high positive value means happiness and 0.0 is the initial value of each unit, negative values express sadness.

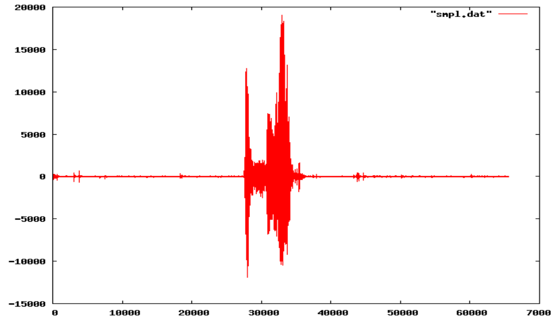
$$F_{t+1}(i) = F_t(i) \pm aC - bD_i \,. \qquad (12)$$

Where $F(i)$ notes the feeling value of unit i on the Feeling Map (zero initially), C notes the continue times of reward (+ a) or punishment(- a), D_i is the Euclidean distance (squared error) between the unit on Feature Map corresponding to i and the input data, a, b are constants, $0 < a < 1$, $0 < b << 1$ .An example of preprocessing of a voice command.

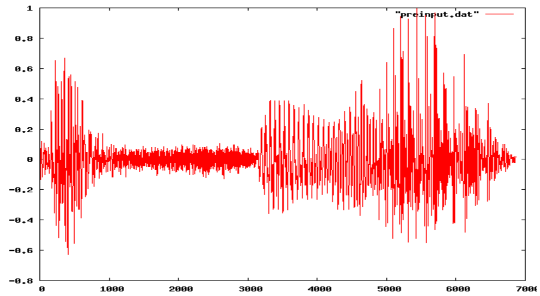
C. *Robot AIBO*

Though the instruction learning system mentioned in A and B sections has been confirmed its effectiveness in simulation experiments in [1], its practicality is not investigated by real robots. In this research, we adopt the system to a pet robot AIBO which is a product of Sony Ltd. manufactured in 2003. The instructions given in Japanese are short and they are expected to be acted by the robot according to its pre-installed standard behaviors, so "instruction" is instead by "command" in the real robot system of this research.

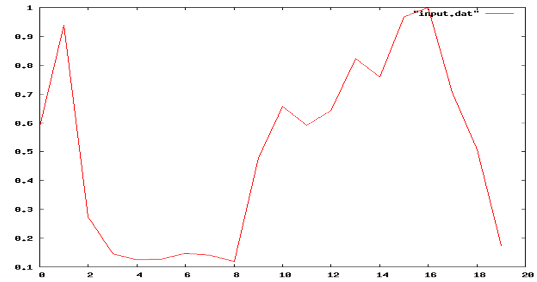
Figure 2 shows the robot AIBO (Sony Ltd., ERS-7, 2003) used in the experiments and Table I shows its specifications. Voice signals are accepted by stereo "Microphones" in the ears when "Record switch" is pushed. "Touch sensor 1 (under the chin)" can be used to send a motion instruction [6], however, it is not used in this study. "Touch sensor 2" on the back can be used to put a "punishment" to the robot, and Touch sensor 3" on the head can be used as a positive reward input.



(a) Sound wave of a command "osuwari"



(b) Regularization result



(c) Normalized feature in 20-demention input space of PL-G-SOM

Figure 3. An example of preprocessing of a voice command.

III. Experiment and Result

The experiment of the voice command learning system described by last section was carried out with robot AIBO. The voice instruction signal was preprocessed by regularization and normalization. The former cut area of standby, and translated the amplitude of sound wave into (-1.0, 1.0) scale. The later calculated the average amplitude of areas which were 20 partitions of the regularized wave. Figure 3 shows these preprocessing results of an instruction in Japanese "osuwari" (sit down).

A. *Experiment Description*

There were 4 kinds of commands spoken in Japanese which instruct AIBO to sit down ("osuwari"), lay down ("fuse"), stand up ("tate") and walk "aruke". 10 samples for each command, i.e., 40 samples were used to train the robot. In a quiet room, not a special acoustic experiment space, a young man in twenties operated the learning system and the procedure is as follows.

Step 1. Push "Record switch" on the back to start recording during 4 second;

Step 2. Preprocessing of voice data, PL-G-SOM learning and categorizing act automatically;

Step 3. AIBO acts a standard motion according to the input voice signal;

Step 4. Instructor push the "Touch sensor 2" on the back to "punish" a wrong action or push "Touch sensor 3" on the head to "award" a correct action of AIBO;

Step 5. LED light bulbs in the face shine to express the feeling of AIBO: "red" face expresses "sadness" which means it failed to recognize a command, "blue" face is a neutral state and "white" shows that a command is recognized and performed correctly.

In practice, "push switch then record" process in Step 1 is not convenient, however, it is a hard work to segment a speech signal online to a series of commands, we are considering to applying a multiple layer SOM to solve the problem in the future. Parameters of PL-G-SOM use the same values given by [1] and iteration steps for each command were 30.

B. *Results and Analysis*

The initial size of PL-G-SOM is 5x5 (25 nodes) and final size was 6x8 (48 nodes). Figure 4 shows 4 constant scenes of an AIBO executed a voice command "sit down" ("osuwari") after training. The change of robot's feeling value during learning process is shown by Figure 5. Each command reached highest value 1.0 of feeling in several iterations which

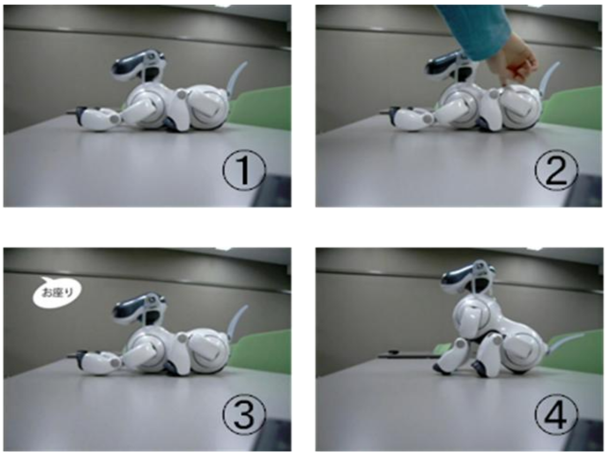


Figure 4. An experiment of PL-G-SOM voice command recognition system using robot AIBO.

means 100% recognition rate and returned to zero which is a neutral state (without any evaluation). Negative values showed the failures of robot actions. Table II shows the success execution rate of each command and results using real robot arrived a middle level between 10% noised simulations and 20% noised.Except the "segmentation problem" mentioned by last section, time lags of motion were also observed during the experiment, i.e., 13 seconds were spent to act after the voice signal generated. This problem is caused by the limitation of robot performance, and is expected to be solved.

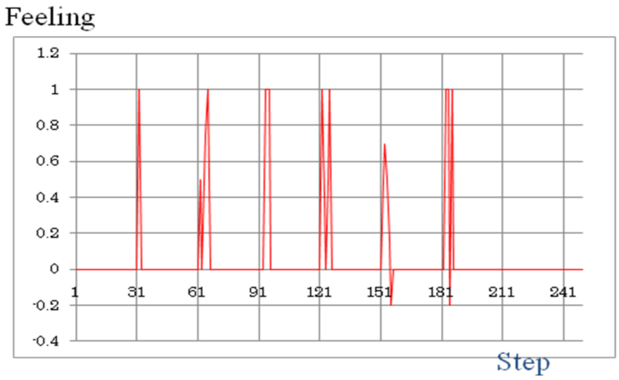


Figure 5. Change of robot's feeling during learning iterations.

TABLE II. COMPARISON OF COMMAND EXECUTION RATE

Voice command in Japanese (action)	osuwa ri (sit)	fuse (lay down)	tate (stand up)	aruke (walk)	Average (4 actions)
Real robot	90.0	80.0	70.0	80.0	80.0
Simulation with 10% noise [1]	87.2	90.8	86.4	82.4	86.7
Simulation with 20% noise [1]	62.0	67.2	48.4	52.8	57.6

IV CONCLUSIONS

A voice command learning system using PL-G-SOM was applied to a real partner robot AIBO. Four kinds of voice commands were learned by 40 samples in Japanese, 80% succeeded on average. The practicability of our system was confirmed by the solid experiment results, and additional learning by PL-G-SOM remains to be applied in the future.

ACKNOWLEDGMENT

We would like to thank T. Hano and T. Komoto for their early work. A part of this work was supported by JSPS-KAKENHI (No. 20500207 and No. 20500277).

REFERENCES

[1] T. Kuremoto, T. Komoto, K. Kobayashi, and M. Obayashi, "Parameterless-Growing-SOM and its application to a voice instruction learning system", J. Robotics, 2010, 9 pages (online).

[2] Bibliography of SOM: http://www.cis.hut.fi/nnrc/refs/

[3] T. Kuremoto, T. Komoto, K. Kobayashi, and M. Obayashi, "A Voice instruction learning system using PL-T-SOM", Proc. Int. Conf. Image and Sign. Proc. (CISP'09), 2009, pp. 4294-4299.

[4] AIBO Official Site: http://www.jp.aibo.com/

[5] S. S. Sutton, and A. G. Barto, Reinforcement learning: an instruction, London: The MIT Press, 1998.

[6] T. Kuremoto, T. Hano, K. Kobayashi, and M. Obayashi, "Robot feeling formation based on image feature", Proc. Int. Conf. Control, Automation and Systems (ICCAS), 2007, pp. 758-761.