A Voice Instruction Learning System Using PL-T-SOM

Takashi KUREMOTO, Takato KOMOTO, Kunikazu KOBAYSHI and Masanao OBAYASHI
Graduate School of Science and Engineering, Yamaguchi University
755-8611 Tokiwadai 2-16-1, Ube, Yamaguchi, Japan
{wu, koba, m.obayas}@yamaguchi-u.ac.jp

Abstract

Learning and relearning ability is important to partner robots. From this point of view, here we propose an intelligent learning system with voice instruction recognition and action learning function. Transient-SOM (T-SOM), an advanced self-organizing map proposed by our previous work for hand gesture recognition and memorization is adopted and improved to be Parameter-Less T-SOM (PL-T-SOM) with an annealing plan. Learning and additional learning experiments used instructions in multiple languages including Japanese, English, Chinese and Malaysian showed the effectiveness of our proposed system.

1. Introduction

Robots which work around us in our daily lives can be called "partner robots". Nowadays, pet robots, housework robots, patrol robots of security, etc, appear more and more. Intelligence, a key word for partner robots, attracts not only users of robots but also designers and developers of robotics.

This study addresses learning and relearning ability of a robot, which can be considered as a kind of intelligence of machines. Input information to a machine is supposed as voice instructions and the output of the machine is expected to be correct actions and visual "feelings" of the machine. This process may be simply categorized into pattern recognition study, however, we think it belongs to machine learning field because both instructor (i.e. maybe a human) and learner (a robot, or a machine) have their subjectivities during the recognition and learning process. From this standpoint, we propose a voice instruction learning system as an internal model of partner robots in this paper.

To classify input information, Kohonen's self-organizing map (SOM) is adopted [1-4]. SOM is a kind of powerful neural network which has been developed to be a number of kinds and applied on pattern recognition, image processing, visualization, intelligent control, and so on [5-14, 18]. It uses a competitive learning process to cluster input features onto a visualized and topological map. Cottrell *et al:* proved SOM's theoretical properties in the general case [5]. More than 7,000 research papers, including its variations and applications, are collected in [6]. However, there are still some problems exist in the standard SOM:

1) The size and the topology of map need to be defined prospectively;

2) How to detect the boundary of clusters;

3) How to tune parameters such as learning rate and neighborhood size.

To solve the first and the second problems, growing structure methods are proposed by [7, 8]. We proposed "Transient-SOM (T-SOM)" [10, 11] which has a hierarchical structure with multiple maps including a memory layer to store "matured unit" and a labeling layer to define the output of feature map which is a standard SOM [12]. T-SOM was applied on a hand image instruction learning system [10, 11] and successfully installed into a pet robot "AIBO" (Sony Ltd., 2003) [13]. In this paper, we intend to improve T-SOM to deal with the third problem of SOM using a new annealing plan proposed recently by Berglund and Sitte [9]. In [9], the parameterless SOM algorithm (PLSOM) calculates learning rate and neighborhood size with local quadratic fitting error (squared error; SE) of the map to the input space. The algorithm is adopted into our T-SOM and a parameter-less T-SOM (PL-T-SOM) is proposed here. Furthermore, T-SOM and PL-T-SOM are applied to construct a voice instruction learning system as an internal model of robot with intelligence of learning and relearning abilities (adaptability) mentioned above.

Four kinds of voice instructions in Japanese were used to be learned at first, and same instructions in English, Chinese and Malaysian were learned additionally in experiments. The results of experiments

978-1-4244-4131-0/09/$25.00 ©2009 IEEE

showed the proposed learning system is effective and PL-T-SOM showed higher efficiency comparing with T-SOM.

2. System

2.1 Structure and Process

The structure (information process chart) of a hand gesture instruction learning system is shown in Figure 1. There are 5 layers including Input Layer, Feature Map, Action Map, Feeling Map, and Memory Map. The flow chart of learning process is shown in Figure 2. Details of processing are described in Subsection 2.2.

2.2 T-SOM and PL-T-SOM

To realize learning, relearning, and additional learning, we proposed Transient-SOM (T-SOM) in [10]. T-SOM consists of 4 layers as same as shown in Figure 1: Input Layer, Feature Map, Action Map and Memory Layer. The standard SOM algorithm is used in T-SOM, but there are two different processes between them: 1) to deal with the oversize of input categories, a long-term-memory layer is added to store matured best match units (*BMU*s). Matching cost is also reduced because the input pattern is compared with the Memory Layer at first, and if it has been learned (stored) then an action can be selected directly without Feature Map; 2) to label categories of input, T-SOM using a reinforcement learning algorithm while learning vector quantizer (LVQ-I) [4] is usually used for the standard SOM. These two processes make Feature Map be no more than a transient process layer, and the units of it is able to be initialized, so we called our SOM as Transient-SOM (T-SOM).

To solve the third problem of standard SOM, i.e., how to reduce learning rate and neighborhood size according to iteration times during training, Berglund and Sitte [9] proposed a parameterless SOM (PLSOM) which used local quadratic error (the distance from the input to the weight vector of *BMU*) to overcome annealing problem of the standard SOM. It is obviously that PLSOM is also available to be adopted into T-SOM. The details are described in Step 3 of T-SOM and PL-T-SOM algorithms later.

T-SOM and PL-T-SOM algorithms are summarized as following:

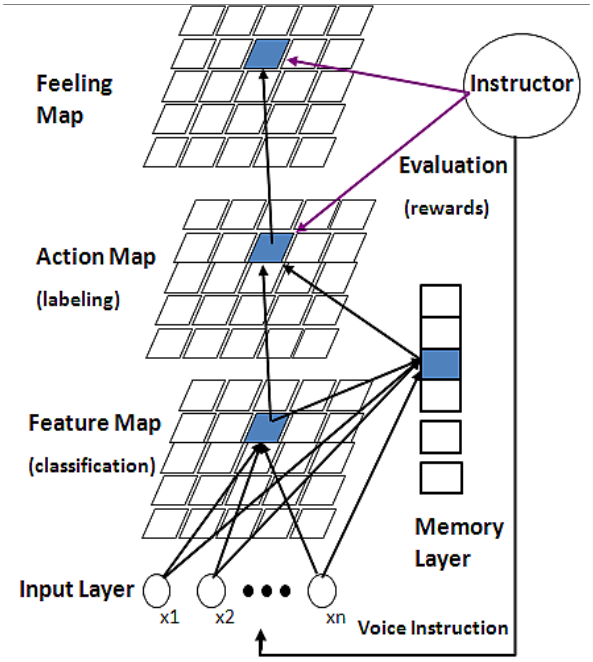


Figure 1. The structure of a voice instruction learning system. using T-SOM and PL-T-SOM.

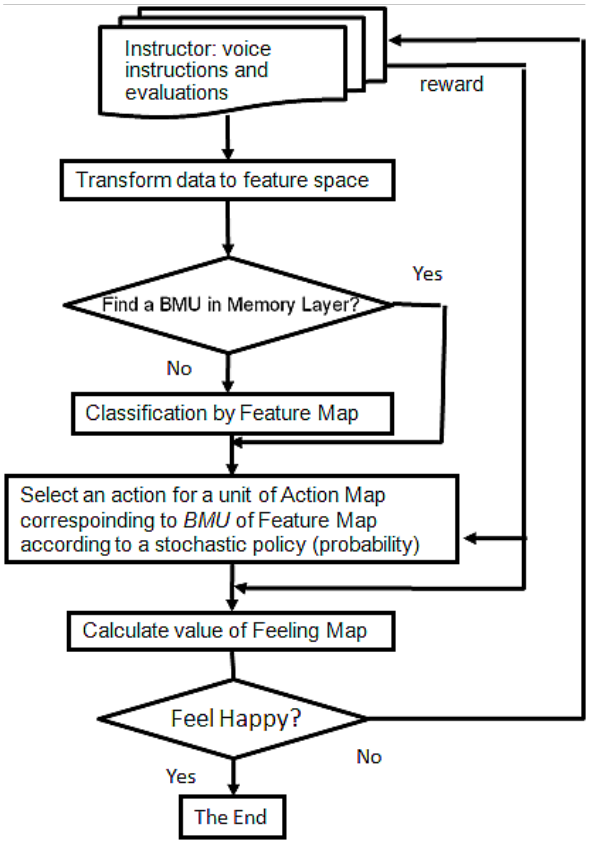


Figure 2. Flow chart of the proposed voice instruction learning system processing.

– Step 1: Initialization. Choose random values (0.0, 1.0) for units m_i $(i = 1,2,...,N \times M)$ of a *2-dimension* map corresponding to an *n-dimension* input space.
– Step 2: Input data. Present a training sample x(x_1, x_2 ..., x_n) to input layer.
– Step 3: Find *BMU,* i.e., the best match unit of Memory Layer or Feature Map. A *BMU c* is decided by using minimum Euclidean distance criterion Eq. (1).

$$c = \arg\min_i(\| x - m_i(t)\|) \quad (1)$$

where x is input feature vector (x_1, x_2, ... , x_n). If *c* is found from Memory Layer, then Feature Map (Step 3) is skipped.
– Step 3: Competitive learning. Using the learning rule given by Eq. (2) to update value of m_i.

$$m_i(t+1) = m_i(t) + \alpha(t)h_{ci}(t)[x(t) - m_i(t)] \quad (2)$$

where $\alpha(t)$ is learning rate and $h_{ci}(t)$ is a neighborhood function.

In standard SOM, $h_{ci}(t)$ is given by Eq. (3) and Eq. (4) and $\alpha(t)$ calculated by an annealing scheme given by Eq. (5).

$$h_{ci}(t) = e^{\frac{-d(i,c)^2}{\beta(t)^2}} \quad (3)$$

where $d(i, c)$ denotes the Euclidean distance from the unit *i* to *BMU c*.

$$\beta(t+1) = \delta_\beta \beta(t) \quad (4)$$
$$\alpha(t+1) = \delta_\alpha \alpha(t) \quad (5)$$

where $0<\delta_\beta<1$ and $0<\delta_\alpha<1$.

In T-SOM, $\alpha(t)$=1.0, and h_{ci} is calculated by Eq. (6) simply, i.e., *t* is not considered, no annealing scheme.

$$\begin{cases} \text{If } d(i,c)=0 & \text{then} \quad h_{ci} = 0.5 \\ \text{If } 1 \le d(i,c) < 2 & \text{then} \quad h_{ci} = 0.3 \\ \text{If } 2 \le d(i,c) < 3 & \text{then} \quad h_{ci} = 0.1 \\ \text{If } d(i,c) \ge 3 & \text{then} \quad h_{ci} = 0 \end{cases} \quad (6)$$

In PL-T-SOM, $\alpha(t)$, $\beta(t)$, δ_α and δ_β are replace by a parameter *I* which is a predetermined iteration times (Eq. (7) and Eq. (8)).

$$\beta(t+1) = (1 - t/I)\, \beta(t) \quad (7)$$
$$\alpha(t+1) = (1 - t/I)\, \alpha(t) \quad (8)$$

For $\alpha(0)$, $\beta(0)$ can be set initially as 0.8 for example, less parameters is used in PL-T-SOM, and fine modification of Feature Map is able to be realized by this annealing plan.
– Step 4: Vector quantization (labeling). After sufficient iterations of Step 3, i.e., if the distance from a *BMU* to the input is less than a threshold value, the input pattern is classified to be a unit of Action Map. This process is as same as LVQ-I [4], but labeling units of Action Map is executed by a reinforcement learning algorithm described in Step 5.
– Step 5: Action learning. Using a reinforcement learning algorithm described in subsection 2.4, robot learns to select "correct actions" for instructor.
– Step 6: Additional learning. For new instruction learning, T-SOM and PL-T-SOM store the succeeded unit weights into Memory Layer, and reset the units of Feature Map into random value. Additional learning or refresh learning then is able to repeat from Step 1.

2.3 Extract Feature

SOM has been widely used in many pattern recognition fields including speech signal processing [14, 15]. Data of voice is a kind of time series data, and recognition can be executed using many methods such as frequency analysis, hidden Markov models (HMM), and SOM. We tried to use average frequency and average amplitude of sound wave in a series of intervals to extract the feature of a voice instruction, and obtained similar results of recognition accuracy. Left column in Figure 3 shows noise eliminated and normalized voice instruction inputs, and right column in Figure 3 shows features extracted by average amplitude in 20 intervals (dimensions). This method is used in learning and additional learning experiments and the details is given by Section 3.

2.4 Reinforcement learning of action

Reinforcement learning (RL) is a powerful learning algorithm by awarding a learner for correct action (adaptive behavior to its environment), and punishing anti-adaptiveactions, i.e, trial-and-error search, widely used in autonomous system, intelligent control, nonlinear prediction and so on [12]. The elements of the learning system usually include: 1) action policy of agent; 2) reward function (reward or punishment from environment) and 3) value function of state-action. In this research, the instructor presents instructions in voice to robot that means a state of the environment s_t is observed by the robot (learner), the robot intends to select a valuable action $a_t(i)$ according to a stochastic action policy π given by Gibbs distribution

(Boltzmann distribution) as shown as Eq. (9).

$$\pi_t(a_t(i)\,|\,s_t) = \frac{e^{\frac{Q_t(s_t,a_t(i))}{T}}}{\sum_{j\in A} e^{\frac{Q_t(s_t,a_t(i))}{T}}} \qquad (9)$$

where T is a parameter named temperature. Now suppose there are $N \times M$ units exist on Action Map, i.e., $N \times M$ states exist in the environment of Markov decision process (MDP), each unit has 4 Actions to be selected available, then a Q-value table can be established as shown in Table 1. The values in the table corresponding to Unit and Action which is random value at firs express initial value of Q_t in Eq. (9). And when an action is selected according to Eq. (9) and performed by robot, instructor evaluates the action by giving a reward/punishment r to robot. The reward is accepted and used to modify the value of Q_t in Table 1 by Eq. (10).

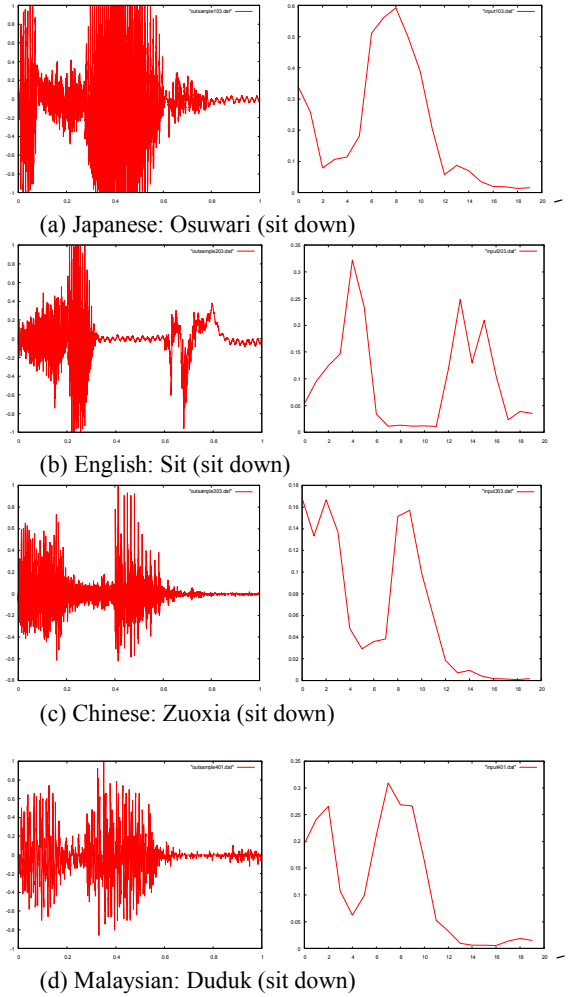
$$Q_{t+1}(s_{t+1}, a_{t+1}(i)) = Q_t(s_t, a_t(i)) + r \qquad (10)$$

Table 1. Q-value table. Each unit of Action Map has a $Q_t(s_t, a_t(i))$ value corresponding to an action.

Unit	Action 1	Action 2	Action 3	Action 4
1	6	2	-8	0
2	10	1	0	1
p	-26	3	5	2
$N \times M$	0	2	7	2

2.5 Feeling Map

To express the degree of how a voice instruction is learned by robot, a Feeling Map which has the same number of units with Action Map is designed (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. The learning algorithm is used as same as [17], also used in [10] and [11] so we do not describe it in detail here.



(a) Japanese: Osuwari (sit down)

(b) English: Sit (sit down)

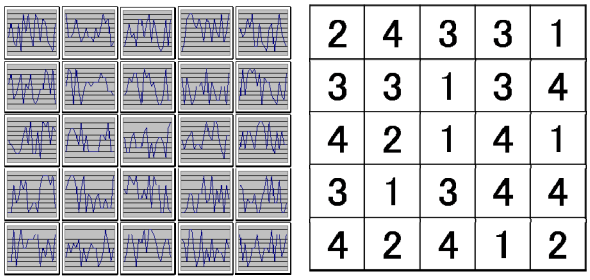
(c) Chinese: Zuoxia (sit down)

(d) Malaysian: Duduk (sit down)

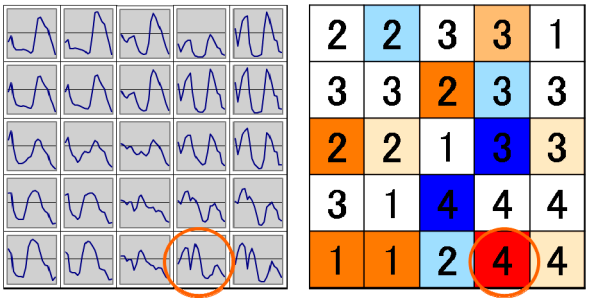
Figure 3. An instruction (sit down) features input to robot in different language. Left: sound waves; right: normalized features.

3. Experiment and Result

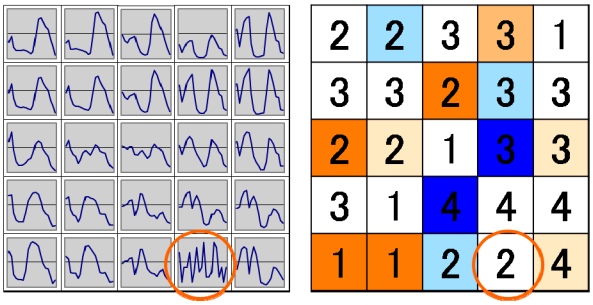
Four kinds of voice instructions were used in experiments: sit down, lie down, stand up and walk. Instructions in Japanese were used to training the system then other 3 languages were used to confirm additional learning ability of the system. Figure 3 shows an example of instruction "sit down" pronounced in Japanese ("Osuwari"), English ("Sit"), Chinese ("Zuoxia") and Malaysian ("Duduk"). The original voices were recorded and executed preprocessing, i.e., noise elimination and normalization.



(a) (0 time) Feature map (left), Action (Feeling) map (right). Radom values were used.



(b) (58 time) Feature map (left), Action (Feeling) map (right). Action 4 matured (finished).



(c) (59 time) Feature map (left), Action (Feeling) map (right). Matured units were initialized.

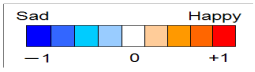


Figure 4. Change of each map during learning process. (a) shows the initial state of units. (b) shows results after 58 learning iterations. (c) shows an operation for additional learning (relearning), and data of the removed unit was stored in memory layer of T-SOM.

There were 3 samples were used for each kind of language while 4 actions with 48 samples. Parameters used in the experiment are shown in Table 2.

Figure 4 shows the change of internal state of Feature Map (left column), Action Map (numbers in right column) and Feeling Map (Depth of color in right column). Random values, random numbers and 0.0 were used initially as shown as in Figure 4 (a).

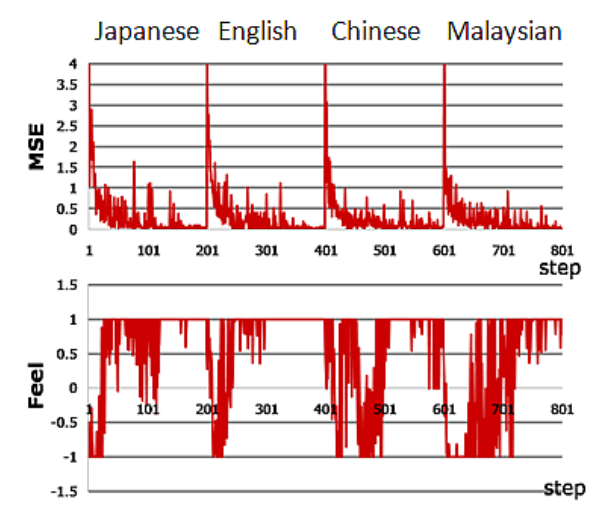


Figure 5. Recognition error of T-SOM reduced with learning (above), while the value of feeling increased with the training steps (under).

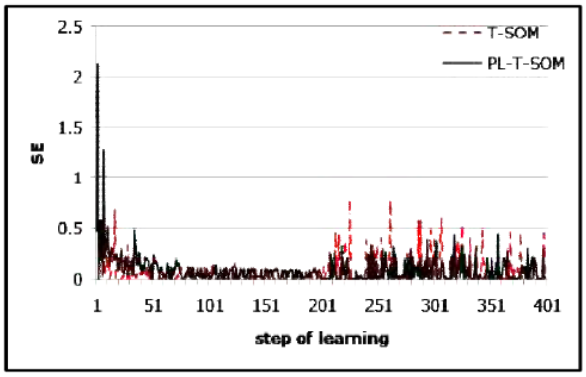


Figure 6. Comparison of recognition error (SE: squared error between input and *BMU*) in learning process between T-SOM and PL-T-SOM.

At the time of iteration 58, Action 4 was learned perfectly because the feeling value rose to the highest value 1.0, which means that the distance from the input to the *BMU* of Feature Map was reduced to a convergence either (Figure 4 (b)). The matured units of each map were initialized again in the next step to prepare for additional learning as shown as in Figure 4 (c).

Figure 5 shows mean squared error (MSE) reduced (upper) and feeling values increased (lower) with learning steps in ten runs of learning and additional experiments using T-SOM.

All instructions were trained till MSE reduced to convergent levels, however, the last two additional learning, i.e., Chinese and Malaysian, showed disturbances of matching errors.

Table 2. Parameters used in the experiment.

Description	Symbol	Quantity
Size of T-SOM	$N \times M$	5×5
Size of PL-T-SOM	$N \times M$	5×5
Iteration times for one instruction	I	200
Temperature	T	0.25
Number of instructions (actions)	$a(i)$	4
Number of additional learning	-	4×3
Reward for one action selected	r	10.0
Parameters of Feeling Map	Ref.[17]	0.2, 0.05
Number of samples	-	48
Sampling rate	-	8KHz
Sample size	-	8bit
channel	-	monaural

The comparison between T-SOM and PL-T-SOM was performed with the total SE (summation of MSE of all instructions in 4 languages), and the results showed PL-T-SOM proposed in this paper obtained better convergence than conventional T-SOM (Figure 6). The reason can be considered that annealing plan used in PL-T-SOM tuned the weights of Feature Map more accurately in the training process.

4. Conclusion

A voice instruction learning system is proposed as an internal model of partner robots in this paper. A novel self-organizing map, PL-T-SOM is developed to recognize and store input patterns of voice instructions. A feeling map is proposed to calculate recognizing rate and degree of action learning in the learning system. Experiment results showed proposed PL-T-SOM had higher learning efficiency than T-SOM. The proposed system is expected to be applied on real partner robots online to confirm its practical performance in the future.

Acknowledgments

This work was supported by Grant-in-Aid for Scientific Research (JSPS No.18500230, No.20500207, No.20500277).

References

[1] Kohonen T., "Self-organized formation of topologically correct feature maps", *Biol. Cybernet.,* 43 (1982), pp. 59-69

[2] Kohonen T., "Analysis of a simple self-organizing process", *Biol. Cybernet.*, 44 (1982), pp. 135-140

[3] Kohonen T., "The self-organizing map", *Neurocomputing*, 21 (1998), pp. 1-6

[4] Kohonen T., *Self-Organizing Maps*, Springer Series in Information Sciences, (1995)

[5] Cottrell M., Fort J.C. and Pages G., "Theoretical aspects of the SOM algorithm", *Neurocomput.*, 21 (1998), pp. 119-138

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

[7] Bauer H.-U. and Villmann Th., "Growing a hypercubical output space in a self-organizing feature map", *IEEE Trans. Neural Networks*, 8-2 (1997), pp. 218-226

[8] Dittenbach M., Merkl D., Rauber A., "The growing hierarchical self-organizing map", *Proc. IEEE Int. Joint Conf. Neural Network (IJCNN '00)*, 6 (2000), pp. 15-19

[9] Berglund E., Sitte J., "The parameter-less self-organizing map algorithm", *IEEE Trans. Neural Network*, 17-2 (2006), pp. 305-316

[10] Kuremoto T., Hano T., Kobayashi K. and Obayashi M., "For Partner Robots: A Hand Instruction Learning System Using Transient-SOM", *Proc. The 2nd Int. Conf. on Natural Computation and The 3rd Int. Conf. on Fuzzy Systems and Knowledge Discovery (ICNC '06-FSKD'06*), (2006), pp.403-414

[11] Hano T., Kuremoto T., Kobayashi K., and Obayashi M., "A Hand Image Instruction Learning System Using Transient-SOM", *Trans. on SICE (Society of Instrument and Control Engineering)*, 43-11 (2007), pp.1004-1006

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

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

[14] Salem Z. N. B., Boougrain L. and Alexandre F., "Spatio-temporal biologically inspired models for clean and noisy speech recognition", *Neurocomput.*, 71 (2007), pp. 131-136

[15] Moschou V., Ververidis D. and Kotropoulos C., "Assessment of self-organizing map variants for clustering with application to redistribution of emotional speech patterns", *Neurocomput.*, 71 (2007), pp. 147-156

[16] Iwasaki H. and Sueda N., "A system of autonomous state space construction with the self-organizing map in reinforcement learning", *Proc. of the 19th Ann. Conf. of the Japanese Society for Artif. Intell.*, (2005), pp. 1-4

[17] Kuremoto T., Hano T., Kobayashi K. and Obayashi M., "Robot feeling formation based on image feature", *Proc. Int. Conf. on Control, Automation and Systems (ICCAS*), (2006), pp.758-761

[18] Aziz Muslim M., Ishikawa M. and Furukawa T., "Task segmentation in a mobile robot by mnSOM: a new approach to training expert modules", Neural Comput. & Applic., 16 (2007), pp. 571-580