

Human Preference Learning by Robot Partners Based on Multi-objective Behavior Coordination

Naoyuki Kubota

Graduate School of System Design,
Tokyo Metropolitan University
6-6 Asahigaoka, Hino-shi, Tokyo,
191-0065, Japan
kubota@tmu.ac.jp

Aiko Yaguchi

Graduate School of System Design,
Tokyo Metropolitan University
6-6 Asahigaoka, Hino-shi, Tokyo,
191-0065, Japan
yaguchi-aiko@sd.tmu.ac.jp

Utaki Ishikawa

Graduate School of System Design,
Tokyo Metropolitan University
6-6 Asahigaoka, Hino-shi, Tokyo,
191-0065, Japan
ishikawa-utaki@sd.tmu.ac.jp

Abstract—This paper discusses human preference learning by robot partners through interaction with a person. We use a robot music player; miuro, and we focus on the music selection for providing the person with comfortable sound field. First, we propose a control architecture of miuro based on autonomous behavior mode, interactive behavior mode, and human control mode. Next, we propose a learning method of the relationship between human position and its corresponding music selection based on Q-learning. Furthermore, we proposed a similarity matrix to reduce the learning time of Q-learning. The experimental results show that the proposed method can learn the relationship between human position and its corresponding human preferable music.

Keywords— Robot Partners; Sound Field; Q-Learning; Boltzmann Selection; Fuzzy Control; Multi-Objective Behavior Coordination

I. INTRODUCTION

Recently, various types of human-friendly robots such as amusement robots, robot pets, and robot partners have been developed until now [1-6]. The advantage of robots is to realize human- or animal-like interaction and communication with people. A robot partner will be a concierge in the information service through interaction and communication. We can have three different types of robot partners from the interactive point of view (Fig.1) [7]. One is a physical robot partner. We can interact with the physical robot partner by using multi-modal communication like a human. The next one is a virtual robot partner. The virtual robot partner is in the virtual space built in the computer, but we can interact with it through the virtual person or robot in the virtual space. The last one is a pocket robot partner, and this is considered as the overlapping subcategory of physical and virtual robot partners. The pocket robot partner has a physical body but no mobile mechanism. We can easily bring it everywhere and can interact with the robot partner by touch and physical interface. The interaction style of these three robots is different, but they share the same personal database and interaction logs, and can interact with the person based on the same interaction rules independent from the style of interfaces. The person can accept the information service ubiquitously from a change of the style of robot partners according to the situation.

In general, the basic behaviors for robots are designed beforehand, but robot partners should perform social learning through interaction and communication with people. Various types of human-friendly robots have been developed, but the aim of most previous works is in the interaction between a robot and person, not the behavioral learning. Such a robot is used for direct interaction and communication with a person, but the mood and comfort by co-existing with the robot have not been discussed so much. Therefore, we discuss the mood and comfort for a human by co-existing with a robot partner. Sound field is one of the most important roles in constructing comfortable space. We proposed the learning method of human preferable music selection based on Boltzmann selection, and showed the experimental results of the relationship learning through continuous interaction [8]. Furthermore, we showed that the learning results by the pocket robot partner could be seamlessly reflected to the interaction rules of the physical robot partner. However, we do not discuss the location- or situation-dependent learning in the learning to extract human preference. Therefore, we propose a Q-learning algorithm for acquiring the relationship between human position in a room and its corresponding preferable music selection to construct a comfortable space by a robot partner. Furthermore, we should consider the regulation of the learning speed through continuous interaction in order to reduce the human load and the times of human interactions. Therefore, we propose a learning method based on similarity matrix used for evaluating human preference according to the similarity among songs.

This paper is organized as follows. Section 2 explains the concept of informationally structured space [8,9] used for the control and information processing of robot partners through wireless network. Next, we explain the hardware of robot partners used in this study, and the control architecture of the robot partners. Section 3 proposes decision making and learning method of robot partners based on computational intelligence. Section 4 shows several experimental results of the proposed method. Section 5 discusses the effectiveness of the proposed method, and concludes this paper.

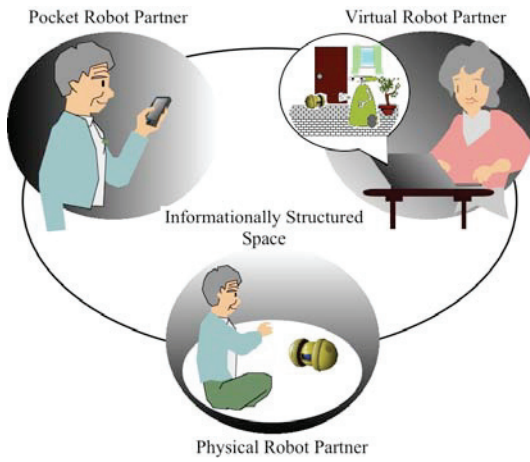


Fig.1. Interaction with robot partners.

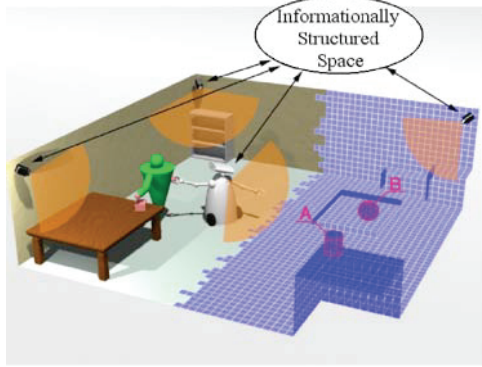


Fig.2. The concept of informationally structured space.

II. INFORMATIONALLY STRUCTURED SPACE FOR ROBOT PARTNERS

A. Informationally structured space and information service

Basically, information service to a person is used to find the required information from huge size of data, and provide them to the person in the suitable form. Robot partners are used as a concierge of daily information service, and require the emerging synthesis of information technology (IT), network technology (NT), and robot technology (RT). NT can provide the robot with computational capabilities based on various types of information outside of robots. Actually the robot directly receives the environmental information through a local area network without the measurement by the robot itself [10]. Wireless sensor networks realize to gather the huge data on environments [11]. However, it is very difficult to store all of huge data in real time. Furthermore, some features should be extracted from the gathered data to obtain the required information. Therefore, the environment surrounding people and robots should have a structured platform for gathering, storing, transforming, and providing information. Such an environment is called informationally structured space (Fig.2).

Basically, information is defined as meaningful or valuable data from the viewpoint of information science. On the other

hand, sensory information includes visual information, audio information, tactile information, and olfactory information, from psychology [12-14]. This means we can perceive the meaning or value from such sensory information to perceive the environment. In this study, we try to realize information service by music, sound and voice. Especially, we discuss the method of providing a person with the comfortable mood by the sound field using music suitable to the situation.

B. A robot partner miuro

We use a robot music player called miuro as a physical robot partner developed by ZMP Inc. in Japan [4]. This robot can act in autonomous motion with the big size of wheels attached with both sides (Fig.3). The robot has many sensors such as a camera, 10 infrared distance sensors, touch sensors, and accelerometer. Furthermore, miuro can be controlled by infrared remote controller. The user can play music contents registered beforehand in miuro by the remote controller, and iPod can be connected and built within the body of miuro. The robot also can play music contents in a PC server and programs of Net Radio connected through wireless LAN (802.11b/g) by its high quality speakers.

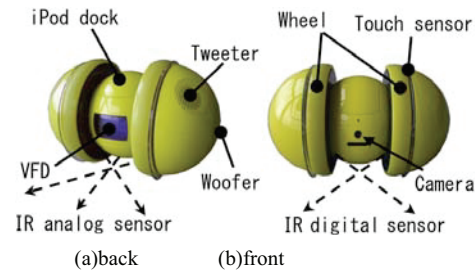


Fig.3. Robot partner; miuro.

Figure 4 shows the roles of miuro in information service. The most important advantage of miuro is in the high quality of music play with autonomous sensing and moving capability. Therefore, miuro can provide a person with the sound field according to the situation extracted as a result of sensing. The second advantage of miuro is in wireless communication with PC server and Internet. The extracted information on preference and custom is stored in the PC server and referred from the PC server. Furthermore, miuro can download music contents from web sites according to the extracted preference of the person, and play them. However, the downloaded music contents might be different from the preference, but sometimes this can be a new finding of music for the person. This means the person does not perform the search of music contents, but miuro can be a concierge of information service.

The roles of music plays can be divided into music appreciation and background music. A person directly enjoys listening music in the music appreciation, while the person enjoys other tasks with listening music in the mode of background music. In the latter, the background music plays the important role in entraining the person into the target task. Based on the portability and mobility of miuro, we focus on the music selection as the background music.

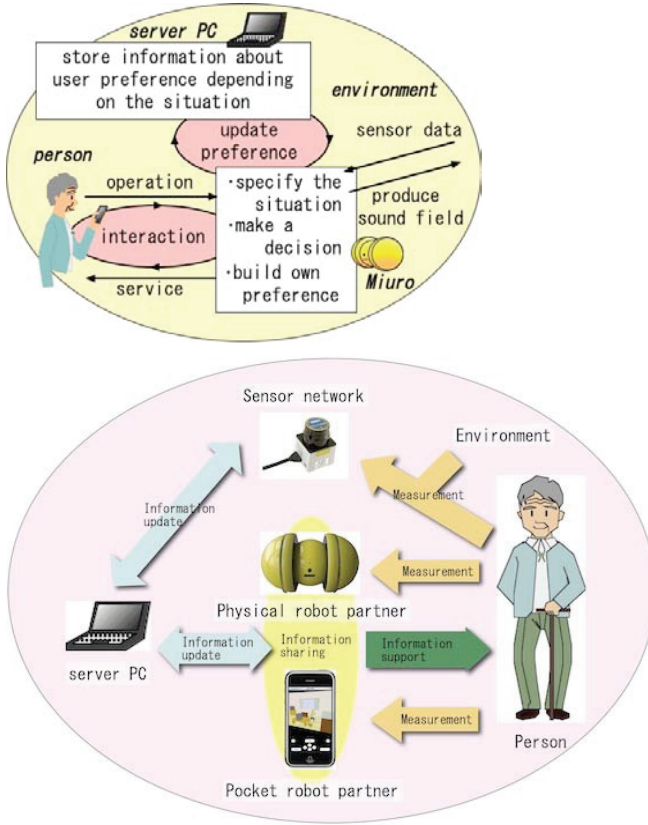


Fig.4. The roles of Miuro in information service.

C. Control architecture of Miuro

Basically, miuro has three behavior modes; human control mode, autonomous behavior mode, and interactive behavior mode. These modes are changed according to the remote controller or human interaction. In the human control mode, the user can directly control miuro, choose the predefined behaviors, and play the music contents by the remote controller. The autonomous behavior mode is composed of random walk, target tracing, information gathering, and map building. The action output of miuro is decided by the multi-objective behavior coordination according to the internal state and local sensing information from external environments. The interactive behavior mode is composed of music play, human tracking, and amusing play.

Figure 5 shows the detail of decision making of miuro. After the sensing, the behavior mode is transited to the human control mode if the remote controller is activated (Fig.5 (1)). In this mode, miuro is directly controlled and used as a music player. On the other hand, if the human interaction is extracted or recognized, the behavior mode is transited to the interactive behavior mode (Fig.5 (3)). In this mode, miuro also aggressively tries to contact and interact with the person, to collect and extract personal information such as preference and custom, and to provide various types of information suitable to the situation.

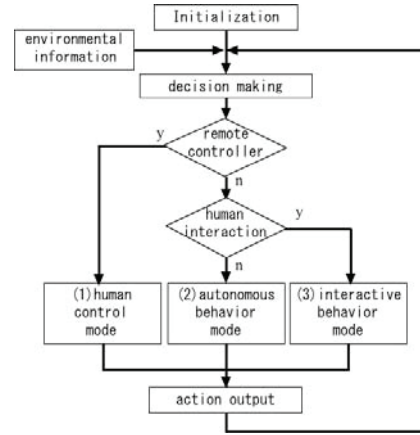


Fig.5. The total procedure of decision making of Miuro.

III. COMPUTATIONAL INTELLIGENCE FOR DECISION MAKING OF ROBOT PARTNERS

A. Autonomous behavior mode

Various intelligent methods for autonomous behavior control have been proposed such as production rules, Bayesian networks, neural networks, fuzzy inference systems, and classifier systems [15-20]. We have applied fuzzy inference systems to represent behavior rules of robots [10,21], because the behavioral rules can be designed easily and intuitively by human behavioral representations. We apply simplified fuzzy inference to the motion control of miuro. In general, a fuzzy if-then rule is described as follows,

If x_1 is $A_{i,1}$ and ... and x_M is $A_{i,M}$
Then y_1 is $w_{i,1}$ and ... and y_N is $w_{i,N}$

where $A_{i,j}$ and $w_{i,k}$ are the Gaussian membership function for the j th input and the singleton for the k th output of the i th rule; M and N are the numbers of inputs and outputs, respectively. Fuzzy inference is performed by,

$$\mu_{A_{i,j}}(x_j) = \exp\left(-\frac{(x_j - a_{i,j})^2}{b_{i,j}^2}\right) \quad (1)$$

$$\mu_i = \prod_{j=1}^M \mu_{A_{i,j}}(x_j) \quad (2)$$

$$y_k = \frac{\sum_{i=1}^R \mu_i w_{i,k}}{\sum_{i=1}^R \mu_i} \quad (3)$$

where $a_{i,j}$ and $b_{i,j}$ are the central value and the width of the membership function $A_{i,j}$; R is the number of rules. Moreover, the k th output y_k is motor output. In general, a mobile robot has a set of behaviors for achieving various objectives, and must integrate these behaviors according to the environmental conditions. Therefore, we proposed the method for multi-objective behavior coordination [22]. The multi-objective

behavior coordination can integrate outputs of several behaviors according to the time-series of perceptual information. A behavior weight is assigned to each behavior. Miuro can achieve a smooth behavior by dynamic update of the behavior weights along perceptual information. According to the behavior weights, the final output is calculated by,

$$y_k = \frac{\sum_{j=1}^K wgt_j(t) \cdot y_{j,k}}{\sum_{j=1}^K wgt_j(t)} \quad (4)$$

where K is the number of behaviors; $wgt_j(t)$ is a behavior weight of the j th behavior over the discrete trials t , and output $y_{j,k}$ is the calculation result from simplified fuzzy inference used for the j th behavior. The update amount of each behavior weight is calculated as follows,

$$\begin{bmatrix} \Delta wgt_1 \\ \Delta wgt_2 \\ \vdots \\ \Delta wgt_K \end{bmatrix} = \begin{bmatrix} dw_{1,1} & dw_{1,2} & \cdots & dw_{1,L} \\ dw_{2,1} & & & \\ \vdots & & \ddots & \vdots \\ dw_{K,1} & & \cdots & dw_{K,L} \end{bmatrix} \begin{bmatrix} ci_1 \\ ci_2 \\ \vdots \\ ci_L \end{bmatrix} \quad (5)$$

where ci_i is the parameter on the perceptual information; L is the number of perceptual inputs.

First, the multi-objective behavior coordination can integrate collision avoidance and target tracing behaviors. The inputs to the fuzzy controller for collision avoidance are the measured distance to the obstacle by infrared distance sensors equipped in front of the robot ($Mc=3$). The inputs to the fuzzy controller for target tracing are the estimated distance to the target point and the relative angle to the target point from the moving direction ($Mt=2$). Outputs are the moving velocity and the steering angle. Final outputs are calculated by the weighted average of outputs from these two behaviors based on the multi-objective behavior coordination according to the degree of danger. The degree of danger is calculated utilizing the minimum value of distances to obstacles by measurement as follows,

$$d = \left(S - \min_j x_j \right) / S \quad (6)$$

where S is the sensing range of infrared distance sensors.

B. Interactive behavior mode

The interactive behavior mode is composed of human tracking, music play, and amusing play. As mentioned before, miuro is designed as a music player, and this paper focuses on the method of providing a person with the comfortable mood by the sound field using music suitable to the situation. The individual preferable sound field corresponding to situations such as cooking, reading, studying, and sleeping is different among people. The robot music player, miuro, should provide the individual person with the comfortable mood, but it is difficult to design the relationship between a situation and its corresponding sound field. Furthermore, such relationship should be extracted and learned through communication and

interaction with the person. In this paper, we assume a specific situation is corresponding to a specific human interaction.

We propose a learning method of the relationship between the human position and its corresponding music play through the interaction with the person. The robot selects human preferable music according to human position. Therefore, the learning results of the robot can reflect human preferences. Based on this assumption, we use Q-learning to obtain the relationship between the human interaction and its corresponding reaction of the robot.

C. Q-Learning for human preference extraction

Reinforcement learning methods build agents that maximize their expected utility only using reward or punishment from the environment [23]. Accordingly, the reinforcement learning does not use explicit teaching signals, but uses evaluative feedback obtained through the interaction with the environment. The tasks to be dealt with in the reinforcement learning are sequential decision problems based on four elements of a policy, a reward function, a value function, and optionally, a model of environment. In this paper, we must consider the continuity of environmental space and the discreteness of songs.

Q-learning is considered as an off-policy TD control method. Watkins proposed a simple method for assigning values to state-action pairs, and called the method incremental dynamic programming. The robot music player of the one-step Q-learning, which is the simplest method, learns the following action-value function,

$$Q(s_t, a_t) \leftarrow (1 - \alpha) \cdot Q(s_t, a_t) + \alpha \cdot \left[R_{t+1} + \eta \cdot \max_{a \in A} Q(s_{t+1}, a) \right] \quad (7)$$

where R_{t+1} is a reward value, α is a learning rate ($0 < \alpha < 1.0$); η is a temporal discounting rate ($0 \leq \eta < 1.0$); and A is a set of songs ($u_1, u_2, \dots, u_S$) corresponding to possible actions; S is the number of songs. It is very difficult to design the reward function for R in this application, because the reward value should be extracted through the interaction with the robot partner. In this paper, we assume a person keeps listening a favorite song. This means the human preference depends on the listening time of each song. If the person changes to another song soon, the preference is not suitable to the state s_t . Therefore, we use the following reward function,

$$R_t \leftarrow \tau(t_p) \cdot R_o \quad (8)$$

where $\tau(t_p)$ is a Sigmoid function; t_p is playing time of song a_t ; R_o is a base value for the reward. Therefore, R_t is high as the playing time is long. Although the update rule is very simple, the action-value function approximates the optimal action-value function if all state-action pairs are updated repeatedly. However, the aim of this paper is not to obtain the optimal action-value function, but to obtain the feasible action-value function to extract human preference. Furthermore, as the difference from previous Q-learning, the state is not directly changed as a result of the selection by a_t . The state is updated

by the human movement in a room. This means we might not need to use Q-learning in this paper, but the advantage of Q-learning is the update of value functions in the multi-stage decision making. In this way, we apply Q-learning because we assume the selection of songs depends on the continuous space in a room.

We apply the softmax action selection (Boltzmann selection), and the i th action a_i is selected with the following probability,

$$p_t(a_i) = \frac{\exp(Q(s_t, a_i)/T)}{\sum_j \exp(Q(s_t, a_j)/T)} \quad (9)$$

where T is the temperature ($T_{\min} \leq T \leq T_{\max}$). When the temperature is high, the robot randomly selects a reaction. As the temperature decreases, the robot deterministically selects the reaction with the high selection strength even if the difference among selection strengths is very small. Generally, T is gradually reduced according to t , but we propose an updating method of the temperature according to the temporally changing pattern of human interactions.

$$T = \begin{cases} T \cdot \lambda & \text{if } x(t) = x(t-1) \\ T / \lambda & \text{otherwise} \end{cases} \quad (10)$$

where λ is the update rate ($0 < \lambda < 1.0$). In this paper, we use music play as reaction of the robot. Therefore, we prepare different types of songs as reaction candidates of the robot.

D. Q-Learning based on Similarity Matrix

There are two important problems in human preference extraction. As the first problem, it is very difficult to evaluate the quality of the extracted human preference. The problem of evaluation has a problem of the mutual nesting. A person can evaluate the preference of music from the candidates of music, while the candidate of music must be selected according to the human preference. This means the quality of the extracted human preference depends on the quality of the selection of music candidates. The second problem is the stability of human preference. This means the human preference can change according to the daily life. In this way, the robot music player should play various types of music, and extract human preference. However, if the robot music player prepares many songs, it takes much time to extract human preference by Q-learning because the state space is large. Therefore, we propose a Q-learning algorithm based on similarity matrix to enhance the learning of Q-learning. The similarity among songs in this paper is not only the objective similarity such as rhythm and tone, but also the subjective similarity like preference among songs.

In the above definition, $Q(v, a_t)$ is discretized where v is the state, and a_t is the song to be played. In previous works, fuzzy Q-learning was proposed to deal with continuous states and actions. In case of fuzz Q-learning, we can deal with the continuity of state and action by the shape of membership functions based on the distance space. However, it is very difficult to represent the distance based on the degree of

similarity among songs by using membership functions. Therefore, we use the similarity matrix in the following;

$$M = \begin{pmatrix} 1 & \cdots & m_{1N} \\ \vdots & \ddots & \vdots \\ m_{N1} & \cdots & 1 \end{pmatrix} \quad (11)$$

where the number of songs is N , and diagonal elements is set at 1; m_{ij} is the degree of similarity between two songs ($0 \leq m_{ij} \leq 1$, $m_{ij} = m_{ji}$). In the proposed method, when the k th song is selected ($a_t = u_k$), Q values are updated as follows;

$$Q(s_t, a_j) \leftarrow (1 - \alpha \cdot \mu_{k,j}) \cdot Q(s_t, a_j) + \alpha \cdot \mu_{k,j} \left[R_{t+1} + \eta \cdot \max_a Q(s_{t+1}, a) \right] \quad (12)$$

(if $\mu_{k,j} > 0$)

where R_{t+1} is a reward value, α is a learning rate ($0 < \alpha < 1.0$); η is a temporal discounting rate; $\mu_{k,j}$ is a similarity value between the k th and j th songs ($0 < \eta < 1.0$); the $\mu_{k,j}$ is calculated in the following;

$$\mu_k = \begin{pmatrix} \mu_{k,1} \\ \vdots \\ \mu_{k,N} \end{pmatrix} = M \mathbf{l} = \begin{pmatrix} 1 & \cdots & m_{1N} \\ \vdots & \ddots & \vdots \\ m_{N1} & \cdots & 1 \end{pmatrix} \begin{pmatrix} l_1 \\ \vdots \\ l_N \end{pmatrix} \quad (13)$$

$$\mathbf{l} = \begin{pmatrix} l_1, \cdots, l_N \end{pmatrix}^T, \quad l_i = \begin{cases} 1 & \text{if } a_t = u_k \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

In the original Q-learning, the Q value of the selected action a_t at the current state s_t is only updated, but the proposed method updates the Q values of other actions not selected at the current state s_t according to the similarity matrix. In this way, the learning time of Q values is reduced.

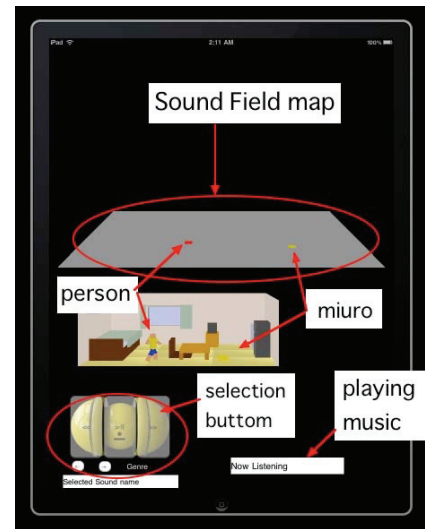
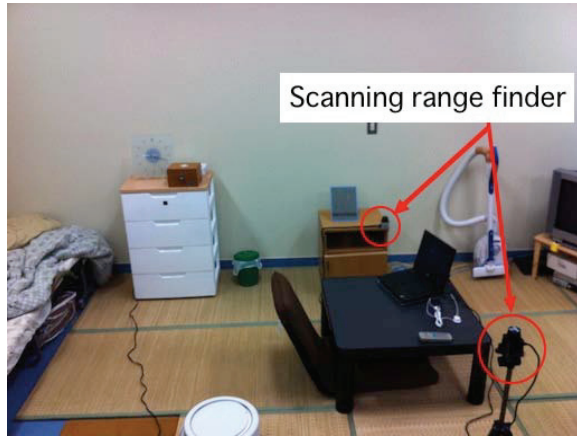


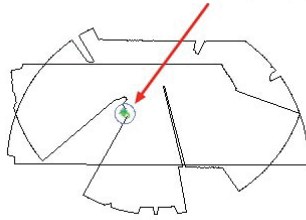
Fig.6. A simulator using iPad

IV. EXPERIMENTAL RESULTS

This section shows several experimental results of the proposed method. In order to realize the efficient and quick learning, we use iPad as the pocket robot partner. The touch interface provides a user with simple and repeatable interaction on the touch screen. After the learning on the pocket robot partner, we can transfer the learning result to the physical robot partner. The human interaction is done with Miuro on the simulator of iPad by the touch interface (Fig.6). The simulator is composed of sound preference map, environmental map simulating to a real room (Fig.7 (a)), and the control buttons of music plays. The human position is estimated by using the measured distance data of two laser range finders (Fig.7 (b)). Figure 8 shows the snapshot of autonomous behavior mode and interactive behavior mode in the pocket robot partner. The robot takes random walk in the autonomous behavior mode in Fig.8 (b), and the robot takes the human tracking behavior in the interactive behavior mode.



(a) One room used for experiments



(b) Human detection by two laser range finders

Fig.7. Experimental environment and human detection with laser range finders.



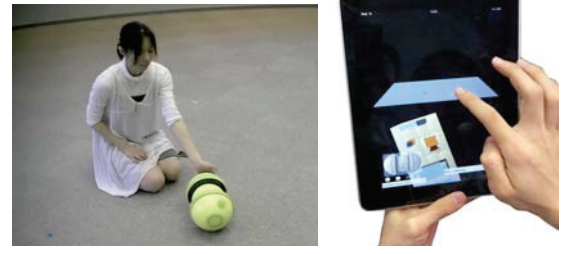
(a) Autonomous behavior mode



(b) Interactive behavior mode

Fig.8. Behavior modes in the pocket robot partner.

We conduct experiments on the learning of the relationship between human location and its corresponding preferable music selection as the generation of comfortable sound field.



(a) with a physical robot partner (b) with a pocket robot partner

Fig.9. Human interaction used in the experiments.

The size of environment corresponding to states is 5×10 . The number of songs is 10 ($S=10$), and the songs are categorized into three different types of classic, rock, and jazz. The current learning states based on the played songs are visualized by using color distribution where classic, rock, and jazz are depicted in red, green, and blue colors. Table 1 shows the values of parameters used in experiments. The values of similarity matrix are designed subjectively. The human position in the simulator is updated according to the estimated position by two laser range finders.

TABLE I. PARAMETERS USED FOR THE LEARNING IN EXPERIMENTS.

Learning rate α	Discount rate η	Standard value of Reward R_0	Number of songs
0.2	0.8	0.5	Classic : 3
			Rock : 3
			Jazz : 4

We conduct an experiment for 30 minutes, and the person selects songs according to the human position in the simulator. The person moves among four points of a bed, chair, cabinet, and sink unit. Figure 10 shows the preference learning in experimental results. Figure 10 (a) shows a result of preference learning at 17 seconds after the start. In this stage, the color distribution is still dark because the learning is not done much. The red color indicates classic songs are played in the sound field map. As time passes, the color of cells is updated gradually. In the final state, the sound preference map is in tinge of red. This means the person listened classic in the bedside, and both classic and jazz in the chair. The learning result is similar with the human preference obtained by the interview after the experiment. In the initial stage of learning, songs are selected almost randomly, but the similar songs based on current learning states are gradually selected by the effect of Q-learning based on similarity matrix. In this way, the proposed method can reduce the learning time while extracting human preference dependent on the human position.

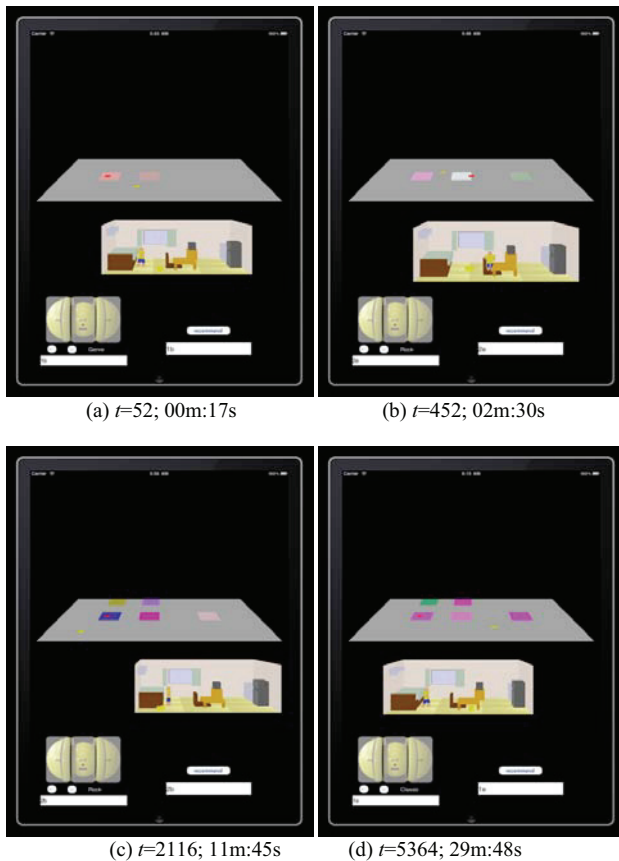


Fig.10. Preference learning in experimental results

V. SUMMARY

This paper proposed a learning method of human preference based on a Q-learning algorithm, Boltzmann selection with temperature control, and similarity matrix. Furthermore, we used a robot music player named miuro as a robot partner. The experimental results show that the proposed method can learn the relationship between human position and its corresponding played songs. Furthermore, the proposed method can provide the person with the preferable song as the comfortable sound field. The essence of the proposed method in this paper is flexibility and adaptability in the update of parameters according to the change of situation and human position. Furthermore, the similarity matrix can reduce the learning time, but the similarity matrix is designed beforehand in this paper. The similarity matrix should be also updated through the interaction with a robot partner.

As future works, we intend to develop the perceptual system of robot partners for specifying the situation sharing with people, and to improve the performance of recognition methods of human gestures and interactions. Furthermore, we will propose a method of replacing song candidates according to human selection results.

REFERENCES

[1] R. C. Arkin, M. Fujita, T. Takagi and R. Hasegawa, An Ethological and Emotional Basis for Human-Robot Interaction, *Robot. Autonom. Syst.* **42** 2003, pp.191–201.

[2] T. Kanda, H. Ishiguro, T. Ono, M. Imai and R. Nakatsu, Development and Evaluation of an Interactive Humanoid Robot Robovie, in *Proc. of IEEE Int. Conf. Robot. Autom.*, Washington, D.C., 2002, pp.1848–1855.

[3] E. Sato, T. Yamaguchi and F. Harashima, Natural Interface Using Pointing Behavior for Human-Robot Gestural Interaction, *IEEE Trans. Ind. Electron.* **54**, 2007, pp.1105–1112.

[4] ZMP Inc. & C., Miuro, The Robot that Play Your Favorite Music at Your Favorite Place (in Japanese), *Journal of Robotics Society of Japan* **26**, 2008, pp.34–35.

[5] N. Kubota, D. Hisajima, F. Kojima and T. Fukuda, Fuzzy and Neural Computing for Communication of A Partner Robot, *Journal of Multiple-Valued Logic and Soft-Computing* **9**, 2003, pp.221–239.

[6] N. Kubota and K. Nishida, Perceptual Control Based on Prediction for Natural Communication of A Partner Robot, *IEEE Trans. Ind. Electron.* **54**, 2007, pp.866–877.

[7] N. Kubota and A. Yorita, Topological Environment Reconstruction in Informationally Structured Space for Pocket Robot Partners, in *Proc. of 2009 IEEE Int. Symp. on Computational Intelligence in Robotics and Automation*, Korea, 2009, pp.165–170.

[8] A. Yaguchi and N. Kubota, The Style of Information Service by Robot Partners, in *Proc. of The 2010 Int. Conf. on Intelligent Robotics and Applications*, 2010.

[9] N. Kubota, T. Obo and B. H. Lee, Localization of Human in Informationally Structured Space Based on Sensor Networks, in *Proc. of IEEE World Congress on Computational Intelligence 2010*, Barcelona, 2010, pp.2215–2221.

[10] N. Kubota and K. Nishida, Cooperative Perceptual Systems for Partner Robots Based on Sensor Network, *Int. J. Comput. Netw. Secur.* **6**, 2006, pp.19–28.

[11] I. Khemapech, I. Duncan and A. Miller, A Survey of Wireless Sensor Networks Technology, in *Proc. the 6th Annual PostGraduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting*, Liverpool, 2005, pp.1–6.

[12] R. L. Gregory, *The Oxford Companion to The Mind*, Oxford University Press, New York, 1998.

[13] M. W. Eysenck, *Psychology: an Integrated Approach*, Prentice-Hall, Addison Wesley Longman Ltd. , Pearson Education, Ltd. , Edinburgh Gate, Harlow, 1998.

[14] K. J. Gergen and M. M. Gergen, *Social Psychology Second Edition*, Springer-Verlag New York, Inc. , New York, 1986.

[15] J. -S. R. Jang, C. -T. Sun and E. Mizutani, *Neuro-Fuzzy and Soft Computing*, Prentice-Hall, Inc. , Upper Saddle River, NJ, 1997.

[16] L. A. Zadeh, eds. G. J. Klir and B. Yuan, *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems*, World Scientific Publishing Co., Inc. , River Edge, NJ, 1996.

[17] J. A. Anderson and E. Rosenfeld, *Neurocomputing*, The MIT Press, Cambridge, MA, 1988.

[18] S. V. Kartalopoulos, *Understanding Neural Networks and Fuzzy Logic*, The IEEE Press, New York, 1996.

[19] S. J. Russell and P. Norvig, *Artificial Intelligence*, Prentice-Hall, Inc. , Englewood Cliffs, NJ, 1995.

[20] H. Liu, A Fuzzy Qualitative Framework for Connecting Robot Qualitative and Quantitative Representations, *IEEE Trans. Fuzzy Syst.* **16**, 2008, pp.1522–1530.

[21] N. Kubota, Multi-Objective Design of Neuro-Fuzzy Controllers for Robot Behavior Coordination, in *Multi-Objective Machine Learning, Studies in Computational Intelligence* **16**, ed. Y. Jin, Springer-Verlag New York, Inc. , Secaucus, NJ, 2006, pp.557–584.

[22] N. Kubota and N. Aizawa, Self-Adaptation in Intelligent Formation Behaviors of Multiple Robots based on Fuzzy Control, in *2009 IEEE International Conference on Fuzzy Systems*, Jeju Island, 2009, pp.900–905.

[23] S. Sutton, G. Barto, *Reinforcement Learning*, A Bradford Book, 1998.