https://doi.org/10.20965/jaciii.2010.p0453

Paper:

A Model for Generating Facial Expressions Using Virtual Emotion Based on Simple Recurrent Network

Yuki Matsui*, **Masayoshi Kanoh****, **Shohei Kato***, **Tsuyoshi Nakamura***, **and Hidenori Itoh***

*Graduate School of Engineering, Nagoya Institute of Technology
Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan
E-mail: {matui, shohey, tnaka, itoh}@juno.ics.nitech.ac.jp
**School of Information Science and Technology, Chukyo University
101 Tokodachi, Kaizu-cho, Toyota 470-0393, Japan
Email: mkanoh@sist.chukyo-u.ac.jp

[Received November 5, 2009; accepted March 23, 2010]

We propose an interactive facial expression model using the Simple Recurrent Network (SRN) for achieving interactions through facial expressions between robots and human beings. The proposed model counts humans in the root system as receivers of facial expressions to achieve a dynamic system bi-directionally affecting humans and robots. Robots typically generate only static changes in facial expression using motion files, so they seem bored, unnatural, and strange to their users. We use interactions between robots and people to diversity the inputs of robots and use emotional state transitions of robots to reduce uniformities in output facial expressions. This paper discusses a dynamic system that causes the proposed model to learn emotional facial expressions based on those of humans. Next, we regard internal states generated by the proposed model as virtual emotions and show that mixed emotions can be expressed by users' inputs from the virtual emotional space. Moreover, based on the results of a questionnaire, we see that facial expressions adopted in the virtual emotional space of the proposed model received high rates of approval from the users.

Keywords: human-robot interaction, simple recurrent network, facial expression, emotion, Ifbot

1. Introduction

Robots are required to have human-like communication channels or interfaces to allow them to live together with people. In addition, they must now have not only physical interactions but also *kansei* for interactions with people. That is, psychological interactions, such as understanding others' emotions, creating harmony, and enjoying communication itself, are necessary in addition to interfaces for the robots to accurately understand instructions given by human beings. In particular, expressions (emotions and affect visible in appearance and gestures) are significant for robots to be amusing. The study focuses on the facial expressions of robots as a medium of communication and as a mechanism to convey emotions.

Fig. 1. Communication between Ifbot and a human subject.

Robots that closely resemble people in appearance, such as Repliee [1] and SAYA [2], are designed to reproduce fine movements that we are normally unconscious of and use Action Units (AU) that formulate facial muscle movements to generate facial expressions similar to those of human beings. Thus, they successfully avoid making unnatural impressions. However, a robot named Ifbot (**Fig. 1**) [3], which we developed, has a cartoonlike facial structure with an emphasis on amusement, unlike that of a human being, so we cannot use AU and some methods to achieve controls to resemble the expressions of human beings. We have therefore discussed neural-network facial expression methods to control facial expressions of Ifbot [4–6]. Gotoh et al. [6] propose a facial expression model for Ifbot in which the emotional space two-dimensionally compressed by an auto-associative neural network is built and used to set weak regions and strong regions of emotions corresponding to subjective human views. An output facial expression is determined by designating coordinate points in the emotional space. Various facial expressions can therefore be generated corresponding to learned facial expressions. However, facial expressions are generated by determining and transiting coordinate points using some emotion control. Our previous researches were limited to a close, static correspondence between emotions and behaviors (facial expressions) in robots. In other words, real people, the receivers of the facial expressions, were excluded from the robot system. We include continuous interactions with users in the system so as to propose a facial expression model of a dynamic system that affects bi-directionally both people and robots.

Human behaviors are diverse, so less-varied, uniform robot facial expressions seem unnatural, strange, and boring to human beings. People perceive environmental stimuli, generate emotional states, and express emotions. The expressions are externally output and change the environment. The next stimulus is then perceived, and a past internal state affects the current emotion forming. Based on people's facial expressions of emotion, the study intends to increase the types of facial expressions from the user's input timing and past internal states to generate various natural facial expressions and to add *kansei* information to interactions with robots. We propose a facial expression generation model using the Simple Recurrent Network (SRN) [7, 8] so as to achieve a dynamic system in facial expressions. Since the SRN is a neural network that gives outputs depending on past state transitions, it is suitable for the facial expression generation of robots. In this paper, we implement a proposed model in a *kansei* robot, Ifbot, and evaluate the interaction between Ifbot and the user to verify the effectiveness of the proposed model as a dynamic system.

This paper includes discussions as follows. Section 2 includes an outline of the facial expression mechanism of Ifbot. Section 3 includes the proposed facial expression model that uses the SRN. Section 4 includes implementation in Ifbot and control. Section 5 includes an example of facial expressions of the proposed model and a qualitative evaluation of facial expressions using virtual emotions. Section 6 includes a quantitative evaluation experiment to compare the proposed model with conventional facial expressions and shows the effectiveness of the proposed model.

2. Facial Expression Mechanism of Ifbot

Ifbot is 45 cm tall and weighs 9.5 kg. It has two arms and moves on wheels. Ifbot can express emotions through cartoonlike, exaggerated, elaborate facial expressions. It has ten motors in its head so that its neck, eyes, and eyelids can move in two axial directions. It uses 104 LEDs in its top of the head, mouth, eyes, cheeks, and tears to generate facial expressions together with color information. **Fig. 2** shows the facial expression mechanism of Ifbot. The facial expression control of Ifbot has the following 15 parameters.

$$S = (\theta_{N1}, \theta_{N2}, \theta_{E1}^{(L)}, \theta_{E2}^{(L)}, \theta_{E1}^{(R)}, \theta_{E2}^{(R)}, \theta_{L1}^{(L)}, \theta_{L2}^{(L)}, \theta_{L1}^{(R)}, \theta_{L2}^{(R)}, L_H, L_M, L_E, L_C, L_T)^T, \quad . \qquad (1)$$

where, $\theta_{\cdot}^{(\cdot)}$ denotes the angle value of the motor and $L_{\cdot}$ denotes the coloration control information value of the LED. These mechanisms allow Ifbot to communicate using rich facial expressions.

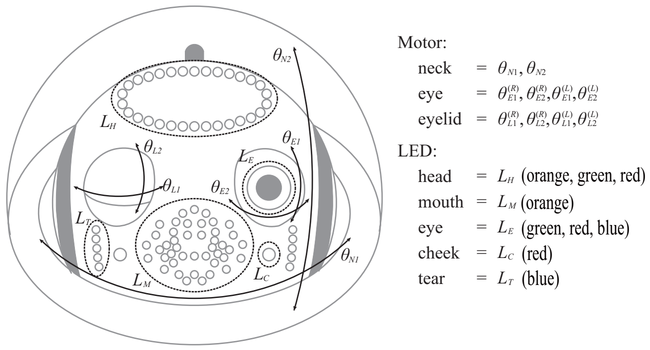


Fig. 2. Expression mechanisms of Ifbot.

3. Interactive Facial Expression Model

A general facial expression of a face robot is enabled by reading an elaborate facial expression file (motion file) corresponding to an emotion when the emotion occurs and by reproducing the file. This is a so-called static facial expression that results in generating the same number of expressions no matter how the user interacts with the robot. Even if the number of motion files is increased, patterned facial expressions may bore the users or make them feel the expressions are strange and unnatural.

This paper therefore proposes an interactive facial expression model in which a robot perceives a stimulus, generates an emotion, and, as a reaction of the stimulus, generates a facial expression. The neural network is useful to handle stimulus and a facial expression control value of the robot, i.e., data that are different in quality [9]. Furthermore, the neural network has a generalization ability that enables flexible responses even to unknown input data. If, like a human, a robot is affected by its current and past psychological states, changes its emotions, and generates facial expressions, temporal changes need to be considered. However, a general multilayer neural network can not handle time series data. We therefore adopt the Simple Recurrent Network (SRN) [7], which gives outputs depending on past state transitions, as a facial expression model.

3.1. Simple Recurrent Network

The SRN (Simple Recurrent Network) (**Fig. 3**) has a structure in which a context layer is added in between an input layer and a hidden layer of a multilayer neural network. The numbers of units of the context layer and the hidden layer are the same, and the units of the hidden layer and those of the context layer correspond one-to-one. In the SRN, the connection weight from the hidden layer to the context layer is fixed to 1, so information of the context layer at a certain period of time reflects information of a past hidden layer. Information stored in the context layer is input again to the hidden layer with the input of the next period of time. Thus, time series data can be processed by adding the context layer, internal feedback.

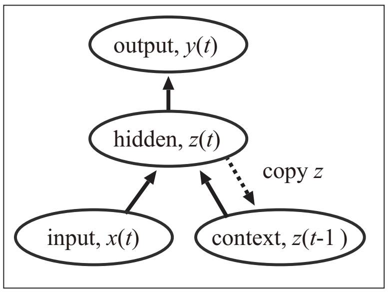


Fig. 3. Simple recurrent network.

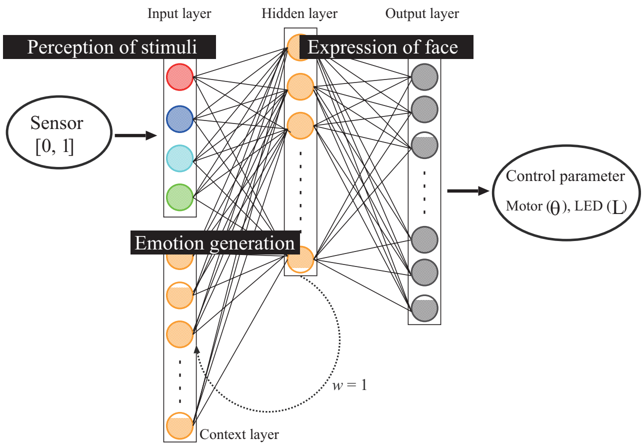


Fig. 4. Proposed interactive facial expression model.

3.2. Facial Expression Model Using SRN

A facial expression model using the SRN is shown in **Fig. 4**. In a normal multilayer neural network, the hidden layer is an internal expression for mapping an input at a certain period of time to an output. On the other hand, assuming that an input is a stimulus and an output is a facial expression, we can consider that in the SRN the hidden layer and the context layer express "emotions," which output a facial expression in consideration of a past stimulus change. Through appropriate learning using the model, it is possible to change the emotions of the robot from past information stored in the context layer and generate facial expressions. Learning is performed in the model by mapping an input stimulus and an output facial expression. That is, to a stimulus, the robot needs to repeatedly learn what behavior (facial expression) to generate. This means a mapping of a stimulus and a facial expression by network learning.

In general, stimuli and facial expressions are mapped by building a mapping relationship between a stimulus and an emotion before connecting the emotion and the facial expression. However, emotional facial expressions are behaviors useful to release or satisfy a sense or a desire and believed to have been acquired by habitual practices. For instance, the self-protection posture of a cat was originally a behavior developed to protect its ears from being bitten by another cat while fighting and thus used in threatening situations "as a habit" [10]. Therefore, the generation of emotions of a robot requires a sort of habit. That is, a robot has to repeatedly learn what behavior (facial expression) to generate for a stimulus. This is to map stimuli and facial expressions through the learning of the network. However, when the learning of the SRN converges, the robot only stores the relation of correspondence between the learned emotions and facial expressions and reacts uniformly. However, after the robot uses the model through interactions with the users, an input timing and an input value to the input layer are different from those while learning. Therefore, the value of the context layer is diversified and the uniformity of the facial expressions output is expected to decrease.

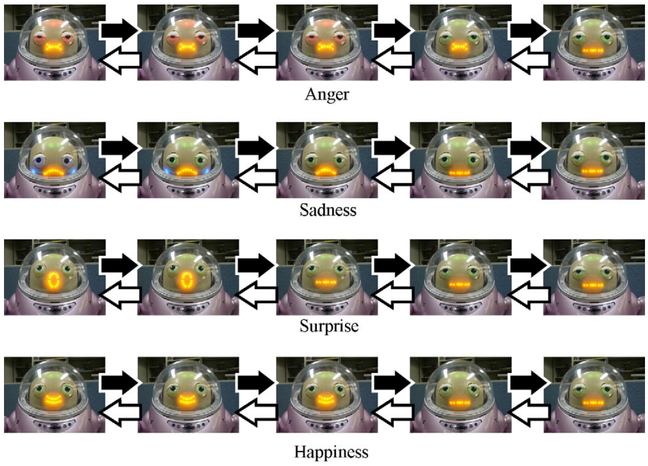


Fig. 5. Teaching data.

In the model, since the network can learn time-series patterns, a facial expression transition when a stimulus is input can be generated based on the internal state change of the network. Also, the generalization ability and the inner feedback of the neural network enable various mixed emotions[1] to be synthesized depending on the input timing of a stimulus. That is, the model uses a pseudo emotion built by the hidden layer and the context layer and based on the learning so as to interpolate the next facial expression from the past emotion change and generate the facial expression. Thus, diversity is expected in its interactions with people.

Figure 5 shows teaching signals used for the learning. For the teaching signals, we used four time-series facial expression data that were prepared in advance: anger, sadness, surprise, and happiness. There are four units in the input layer, 25 in the hidden layer, 25 in the context layer, and 15 in the output layer of the network.

1. Ekman states that facial expressions rarely show a single emotion at any moment and usually show two or more emotions, and he calls them "blended emotions" [11]. Facial expressions indicating "blended emotions" are normally seen on people's faces in daily life and are called "mixed emotions."

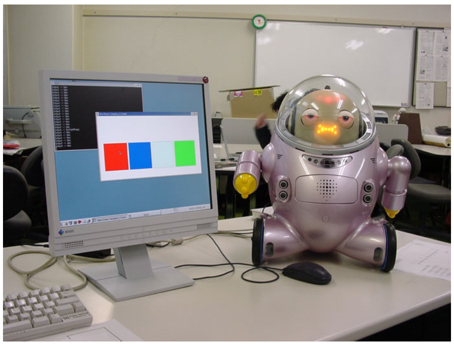


Fig. 6. Evaluation system using Ifbot.

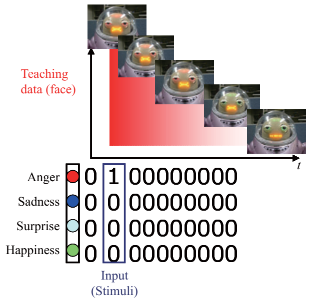


Fig. 7. Learning of *Proposed Robot A*.

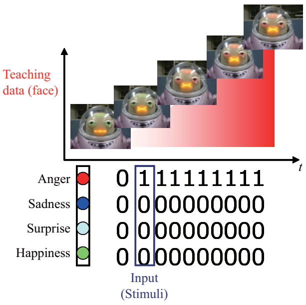


Fig. 8. Learning of *Proposed Robot B*.

4. Building Evaluation System

4.1. Structure of Evaluation System

We built an evaluation system to confirm the effectiveness of the proposed model. **Fig. 6** shows the evaluation system using Ifbot.

In the system, when the user gives the robot a stimulus that generates an emotion at arbitrary timing, the robot generates a facial expression in response to the stimulus. We selected four emotions, anger, sadness, surprise, and happiness, as a stimulus from among six basic emotions [11]. Colored buttons displayed on the monitor in the figure are the input interface that gives the robot a stimulus. The red button corresponds to a stimulus of anger (Anger, *Ang*), the blue to sadness (Sadness, *Sad*), the cyan to surprise (Surprise, *Sup*), and the green to happiness (Happiness, *Hap*). In the experiment, since the facial expression is focused, stimuli and emotions simply correspond one-to-one to make the effect of the proposed model itself clear. The inputs to the system are monitored at intervals of 1 sec. When the user clicks on any of the buttons with the mouse, 1 is input to a unit corresponding to the input layer of the SRN, and, when there is no input, 0 is input. For example, when the anger button is clicked, $(Ang, Sad, Sup, Hap) = (1,0,0,0)$ is input to the network.

Figure 5 shows teaching signals used for the learning. For the teaching signals, we used four time-series facial expression data that were prepared in advance: anger, sadness, surprise, and happiness. There are four units in the input layer, 25 in the hidden layer, 25 in the context layer, and 15 in the output layer of the network.

4.2. Learning of Robot

We believe that the character of the robot can be defined by varying the correspondence of a stimulus and a facial expression through learning. The paper discusses the implementation of two robots (*Proposed Robot A* and *Proposed Robot B*) whose characters are defined by giving different teaching signals. The learning procedure of the proposed model will now be explained.

Proposed Robot A —

Proposed Robot A is a robot which generates an emotional facial expression corresponding to a stimulus shortly after the stimulus is input. That is, the leftmost facial expression is generated when a stimulus is input (**Fig. 5**). **Fig. 7** shows the learning method of *Proposed Robot A*. The figure shows the correspondence between an input (stimulus) and a teaching signal (facial expression) for the learning of an emotional facial expression (anger in the figure). As an input, during time $t = 1$ sec, 1 is given to units of the input layer corresponding to the stimulus, and, during time $t = 2$ to 10 sec, the value of 0 is given as there is no stimulus. In addition, a facial expression change in the direction of the right arrow (**Fig. 5**) is given as the teaching signal so that the robot can learn. As a result, the proposed model can learn to generate an emotional facial expression shortly after a stimulus is input. In the above manner, we caused the robot to learn all four emotions.

Proposed Robot B —

Proposed Robot B is a robot which gradually generates facial expressions of emotion as stimuli are continuously input. **Fig. 8** shows the learning method of *Proposed Robot B*. As an input, 1 is given to units of the input layer corresponding to the stimulus continuously after time $t = 1$ sec, and, a facial expression change in the direction of the left arrow (**Fig. 5**) is given as the teaching signal so that the robot can learn. As a result, the proposed model can learn to generate an emotional facial expression shortly after many stimuli are input. In the above manner, we caused the robot to learn all four emotions.

5. Evaluation of Facial Expressions by Virtual Emotion

The proposed model causes the network to learn the correspondence between a stimulus and a facial expression, and, when the learning is converged, it simply generates only the learned time-series facial expression. However, by inputting unlearned input series coming from the user, the proposed model can generate facial expressions of the robot from the input and the internal state. We have to verify how the internal state of the network changes and generates the facial expressions. Here, as described in Section 3, since the hidden layer corresponds to the internal expression of the facial expression to be output in

Table 1. Test data.

Test data 1	$(\mathscr{A}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N})$
Test data 2	$(\mathscr{A}\,\mathscr{A}\,\mathscr{A}\,\mathscr{A}\,\mathscr{A}\,\mathscr{A}\,\mathscr{A}\,\mathscr{A}\,\mathscr{A}\,\mathscr{A})$
Test data 3	$(\mathscr{A}\,\mathscr{S}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N})$
Test data 4	$(\mathscr{A}\,\mathscr{A}\,\mathscr{A}\,\mathscr{S}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N}\,\mathscr{N})$
Test data 5	$(\mathscr{A}\,\mathscr{A}\,\mathscr{A}\,\mathscr{A}\,\mathscr{A}\,\mathscr{S}\,\mathscr{S}\,\mathscr{S}\,\mathscr{S}\,\mathscr{S})$

$\mathscr{A}$: $(Ang, Sad, Sup, Hap) = (1,0,0,0)$ - Anger.
$\mathscr{S}$: $(Ang, Sad, Sup, Hap) = (0,1,0,0)$ - Sadness.
$\mathscr{N}$: $(Ang, Sad, Sup, Hap) = (0,0,0,0)$ - No stimulus.

the output layer, the hidden layer and the context layer of the proposed model can be regarded as pseudo emotions.

In this section, we analyze the hidden layer, which is the internal state of the proposed model, so that facial expressions are performed according to the learned character definition, and we confirm whether a facial expression with the change in an appropriate emotion is performed when an unlearned input series is input.

5.1. Evaluation Method

We analyze an internal state vector $H(t)$ of the hidden layer units at time t and evaluate $H(t)$ by visualizing it in a two-dimensional space.

First, we assume a set $\mathscr{H}^{Ang} = \{H^{Ang}(t)|t = 0,\ldots,9\}$ of an internal state vector when a teaching signal of anger is input to the learned proposed model. Similarly, we assume $\mathscr{H}^{Sad}$, $\mathscr{H}^{Sup}$ and $\mathscr{H}^{Hap}$ perform a principal component analysis to $\mathscr{H} = \mathscr{H}^{Ang} \cup \mathscr{H}^{Sad} \cup \mathscr{H}^{Sup} \cup \mathscr{H}^{Hap}$ and build a two-dimensional space using the resultant first and second principal components. This is referred to as a "virtual emotional space." The space structure represents the robot emotion characteristics or the robot character definition. Next, the state of the hidden layer when a test pattern is input is mapped in the virtual emotional space so that an internal state change when the unlearned input series is input to the network can be evaluated.

Test data used for the evaluation are shown in **Table 1**. The test data show the inputs in time order. The $\mathscr{A}$ indicates an input to anger, $\mathscr{S}$ indicates an input to sadness, and $\mathscr{N}$ indicates the absence of input. For example, test data 3 indicate that the user inputs sadness right after inputting anger into the robot and inputs nothing afterwards.

5.2. Evaluation Results

Figures 9 and **10** show the virtual emotional space for each robot. They are graphs in which the principal component analysis results of H are plotted. The trajectories of anger, sadness, surprise, and happiness in the figure are graphs of $\mathscr{H}^{Ang}$, $\mathscr{H}^{Sad}$, $\mathscr{H}^{Hap}$ and $\mathscr{H}^{Sup}$ respectively, and they indicate transitions in the internal states of the virtual emotional space. The coordinate point O is of calm (neutral), around which the trajectory is drawn and distributed for each of the emotions.

In the virtual emotional space, the distance from O represents the strength of the emotions. For instance, in the

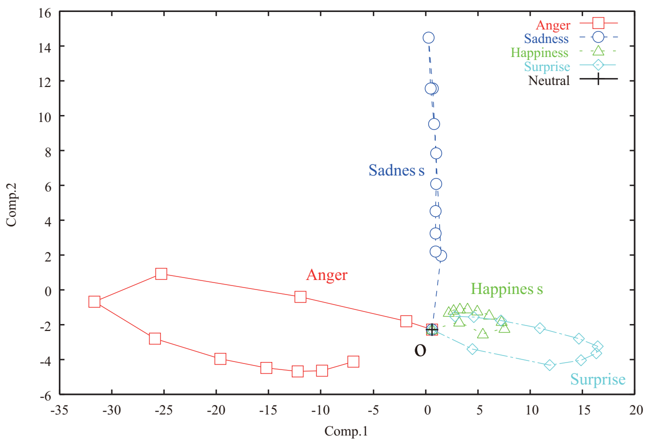


Fig. 9. Virtual emotional space of *Proposed Robot A*.

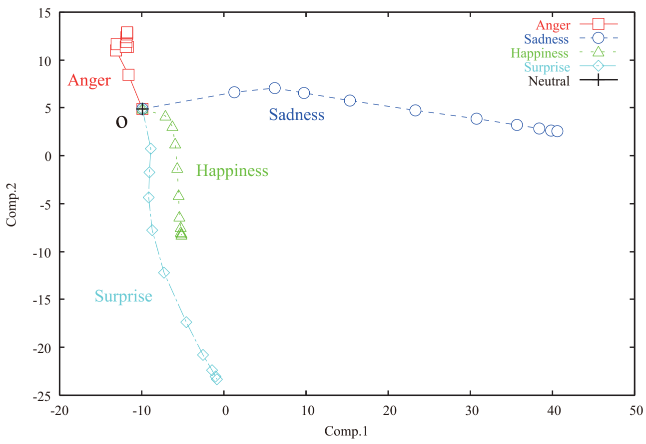
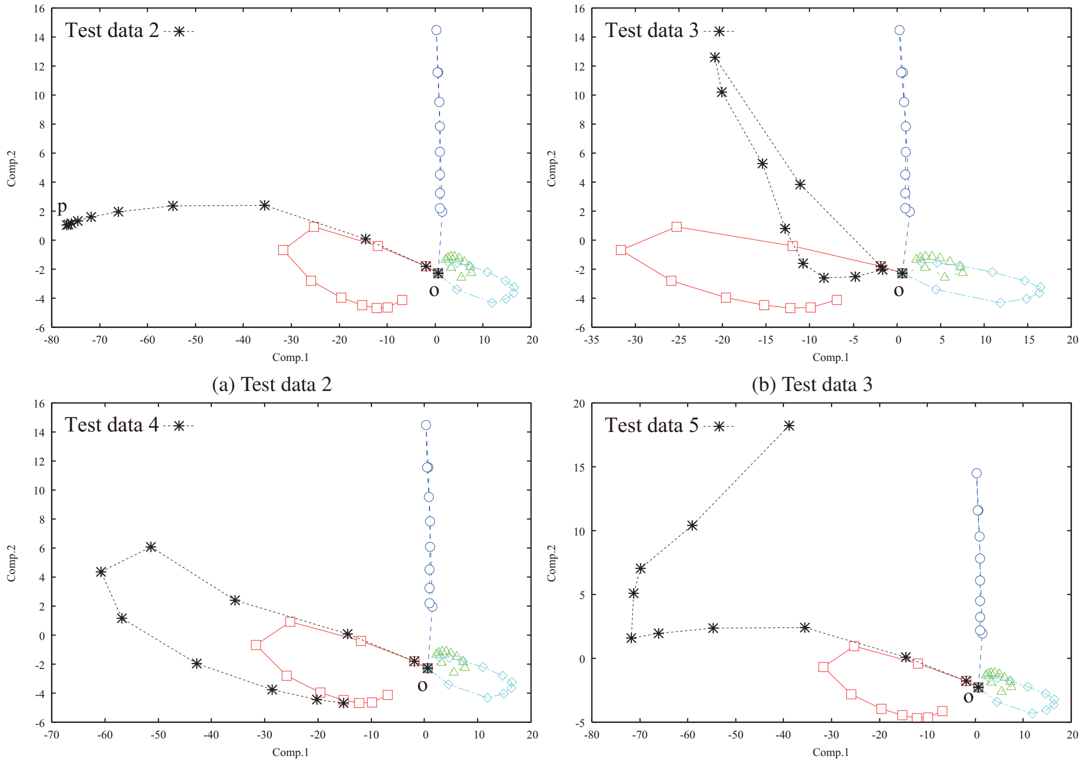


Fig. 10. Virtual emotional space of *Proposed Robot B*.

virtual emotional space of **Fig. 9**, the coordinate point, which constitutes a trajectory right after a stimulus is input, is located farther from O and results in the greatest change of emotion at the farthest coordinate point (a robot facial expression corresponding to each coordinate point is nearly identical to the facial expression transition of the teaching data). Since there is no input after that, the coordinate point returns to near O. Similarly, in the virtual emotional space of **Fig. 10**, the input series when learning is continuously input so that the trajectory radiates from O and results in the greatest change of emotion at the last stimulus input. Thus, the space built acquires a structure that reflects the learning of each robot. Now, we show change in the internal state and corresponding facial expressions when test data is input to each robot and confirm whether it is an appropriate facial expression change according to the learning.

5.2.1. *Proposed Robot A*

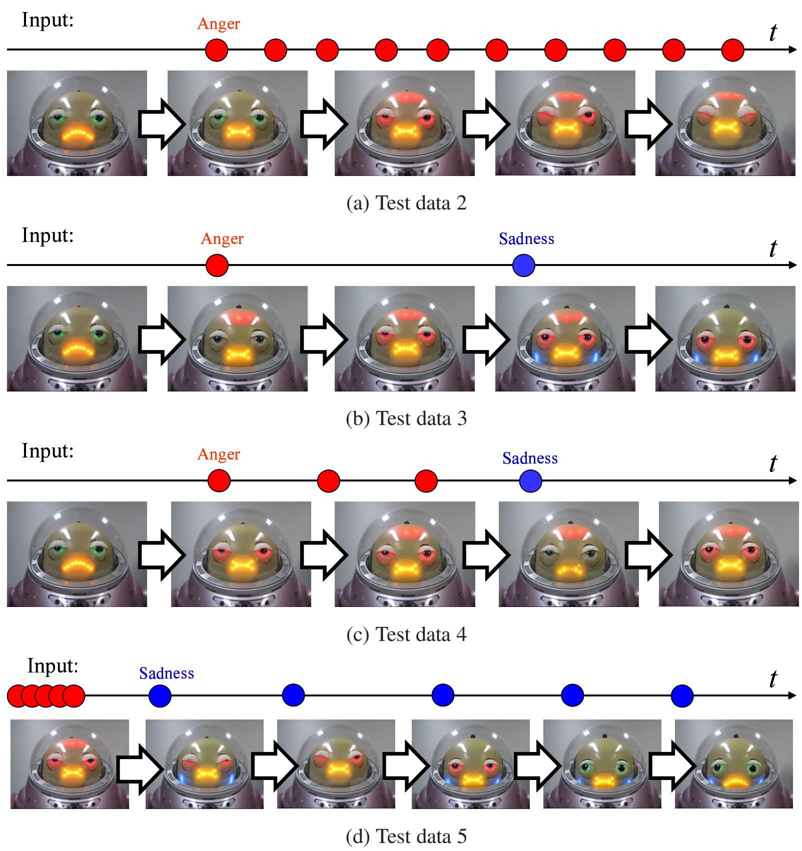
Since test data 1 is of an input series of anger when learning, it draws a trajectory of the anger in the virtual emotional space of **Fig. 9**. **Figs. 11** and **12** show the results and facial expression changes when test data (except test data 1) is input.



(a) Test data 2 (b) Test data 3

(c) Test data 4 (d) Test data 5

Fig. 11. Mapping results in *Proposed Robot A*.



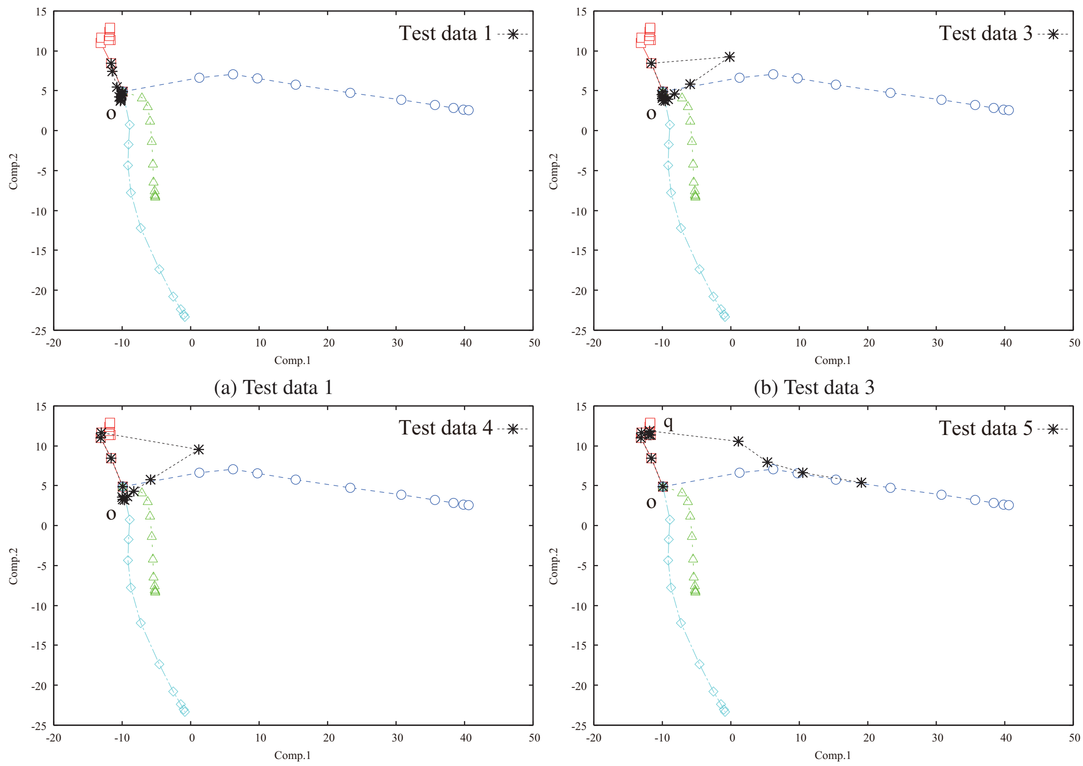
(a) Test data 2

(b) Test data 3

(c) Test data 4

(d) Test data 5

Fig. 12. Facial expressions using *Proposed Robot A*.



(a) Test data 1 (b) Test data 3

(c) Test data 4 (d) Test data 5

Fig. 13. Mapping results in *Proposed Robot B*.

- Test data 2 —

  The trajectory starts from the coordinate point O of calm, passes through the anger region, and shows no change at the coordinate point p (**Fig. 11(a)**). Although at this time the robot generates the facial expression of anger since the stimulus was input, after this point in time it gets angrier (**Fig. 12(a)**).

- Test data 3 —

  $\mathscr{S}$ is input right after $\mathscr{A}$ is input. As a result, in the virtual emotional space, the trajectory passes through the space between the anger region and the sadness region and returns to calm (**Fig. 11(b)**). With the trajectory transition, the robot generates a mixed emotion of anger and sadness (**Fig. 12(b)**). This characteristic facial expression is not included in the teaching signal, and the proposed model can easily generate such a facial expression from the virtual emotional space.

- Test data 4 —

  $\mathscr{S}$ is input once right after $\mathscr{A}$ is input three times. The trajectory drawn in the emotional space does not pass through the sadness region. Instead, it closes in the anger region (**Fig. 11(c)**). The facial expression generated at this time does not include a facial expression of sadness (**Fig. 12(c)**). This indicates that the emotion does not change even if a little sadness is added in a state where the anger has grown due to a multitude of stimuli and reflects the character definition by learning.

- Test data 5 —

  Unlike test data 4, $\mathscr{S}$ is continuously input after $\mathscr{A}$ is continuously input so that the trajectory transits to the sadness region (**Fig. 11(d)**). This indicates that a multitude of sadness stimuli cause the emotion to shift towards sadness even if the internal state of anger has accumulated. At this time, the facial expression gradually changes to the sadness (**Fig. 12(d)**).

5.2.2. *Proposed Robot B*

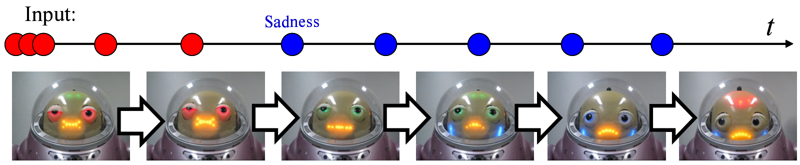
Since the test data 2 is of the input series of anger when learning, it draws a trajectory of the anger in the emotional space of **Fig. 10**. **Figs. 13** and **14** show results and facial expression changes when test data (except test data 2) are input.

- Test data 1 —

  This input series has only one stimulus, and the trajectory of anger is plotted near the coordinate point O of calm (**Fig. 13(a)**). At this time, the facial expression is almost the same as that of calm.

- Test data 3, Test data 4 —

  The trajectory drawn in comparison with the virtual emotional space is small and has a little facial expression change (**Figs. 13(b)** and **(c)**). Unlike *Proposed Robot A*, a mixed emotion is not included in



(d) Test data 5

Fig. 14. Facial expressions using *Proposed Robot B*.

the facial expression that is only generated almost the same as in the case of calm.

- Test data 5 —

  Since at first the anger stimulus is continuously input, the trajectory reached the coordinate points q away from the calm O (**Fig. 13(d)**), and generates the anger facial expression (second frame of **Fig. 14**). After that, the sadness stimulus is input, and the coordinate point moves to the sadness region, so the facial expression of sadness is gradually generated while generating a mixed emotion of anger and sadness (after the second frame of **Fig. 14**).

By varying the teaching signal given to the SRN, we have thus confirmed that the different emotional spaces are built. In addition, without increasing the variation of the facial expression itself, we have thus confirmed that the SRN generated various facial expressions depending on the time series information of inputs (input of the user in interaction). As a result, when the user gives an interaction a dynamic, various facial expressions are generated depending on the psychological state of the robot, a state that varies momentarily.

6. Subjective Evaluation of Facial Expression

The previous section discussed the relationship between an internal state and a facial expression using an input timing and a frequency between two emotions. The present section includes an interaction experiment including the robot and the user in order to quantitatively analyze a transition of an arbitrary emotion triggered by a voluntary interaction of the user. We analyze the result of a subjective evaluation so as to confirm the effectiveness of the proposed method. As a comparison target, we prepared the following facial expression robots (*Control Robot A* and *Control Robot B*) using a conventional, elaborated facial expression.

Control Robot A —

Control Robot A is a robot that generates teaching signals used for the learning of *Proposed Robot A* over time. In other words, a conventional facial expression in which a teaching signal is used as a motion file is performed. More specifically, when one stimulus is input at a certain period of time, a facial expression (the leftmost of the facial expressions in **Fig. 5**) of an emotion corresponding to the stimulus is generated, and then the change in facial expression in the direction of the right arrow is gradually generated. If another stimulus is input in the middle of the facial expression change, a facial expression corresponding to the stimulus is generated.

Control Robot B —

Similar to *Proposed Robot B*, *Control Robot B* gradually performs emotional facial expressions as stimuli are continuously input. More specifically, each facial expression of emotion is designed to transit one by one to a facial expression in the direction of the left arrow in **Fig. 5** as the same stimulus is input using a motion file of a teaching signal of *Proposed Robot B*. If there is no input in the middle, the flow returns to the facial expression of calm.

6.1. Experimental Procedure

Test subjects rank the facial expressions of the four types of robots (*Proposed Robot A*, *Proposed Robot B*, *Control Robot A*, and *Control Robot B*). The robots are presented in a random order for each of the test subjects. The test subjects are allowed to operate the robots to see their facial expressions as long as they like before they are ready to rank the robots. The input/output interfaces are the same in all the robots. We explain to the test subjects in advance that they are expected to evaluate reactions (facial expression generation and change) of the robots depending on the timing of stimuli given by the user. The test subjects were university students of age 20 to 24, including 23 males and five females.

6.2. Evaluation Results

Figure 15 shows a rank-wise frequency distribution of the number of supporting subjects. According to the figure, many test subjects rated the proposed method highest or second highest. After the experiment, many test subjects answered that the proposed method successfully generated interesting facial expressions and felt that it was an elaborate system because mixed emotions were generated. Since the results obtained through the questionnaire were on an ordinal scale, their relation in terms of magnitude is guaranteed. However, the widths of the intervals are unknown and not guaranteed to be equal. Therefore, we calculated a more advanced distance scale from the ordinal scale and used a normalized-rank method [12] for comparison on an interval scale (**Fig. 16**). As a result, we saw a significant difference between the proposed method

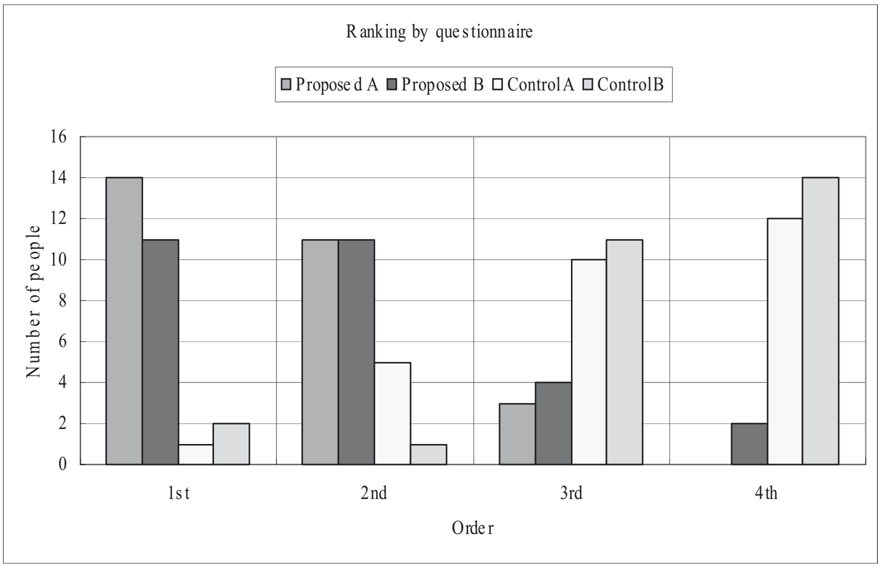


Fig. 15. Frequency distribution of each rank.

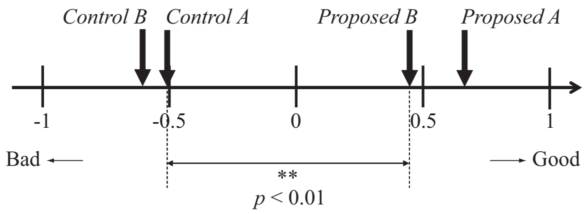


Fig. 16. The distance scale using the normalized-rank approach.

and each of the comparison methods at a significant level of 1%. According to the figure, the proposed method was evaluated more highly regardless of the setting of character definitions.

7. Conclusions

This paper has proposed a facial expression model using the simple recurrent network as a dynamic system in facial expressions including interactions with human beings.

In Section 5, we evaluated the difference in the virtual emotions built depending on the correspondence relation of teaching signals in the learning of the robot. We then confirmed that a facial expression model that can express emotions even to unlearned input is built. In Section 6, we verified the effectiveness of the proposed model through the subjective evaluation of the interaction between a person and Ifbot. As a result, the facial expressions of the proposed model are more interesting than the conventional facial expressions and are evaluated more highly.

However, we have not yet determined whether the psychological effects in the interaction obtained through the experiment described in this paper are because of an increase in variety made possible by the proposed model or because of the simple increase in variety itself. In the future, we have to figure this out by using other methods of adding variety, such as by adding randomness. In the experiment, the evaluation was conducted in a framework of a simplified, short-term interaction to confirm the effectiveness of the proposed model. In a short-term interaction, only the effect of the first impression of the robot and the initial process in which the relationship between a human and a robot are evaluated [13]. As a future issue, the effect of boredom on long-term use has to be researched.

Acknowledgements

Ifbot was jointly developed thanks to the production plan and the supervision by Business Design Laboratory Co., Ltd. In particular, we owe Brother Industries, Ltd. for the shape design and the facial expression control mechanism and Brother Industries, Ltd., ROBOS Co., and Nagoya Institute of Technology for the facial expression control software. We would greatly appreciate those who were involved. A part of the study was conducted with supports from Grant-in-Aid for Young Scientists (A) of the Ministry of Education, Culture, Sports, Science and Technology (A) (# 20680014) and Artificial Intelligence Research Promotion Foundation.

References:

[1] T. Minato, M. Shimada, S. Itakura K. Lee, and H. Ishiguro, "Does Gaze Reveal the Human Likeness of an Android?," Proc. of the 4th IEEE Int. Conf. on Development and Learning, pp. 106-111, 2005.

[2] T. Hashimoto, S. Hiramatsu, T. Tsuji, and H. Kobayashi, "Development of the Face Robot SAYA for Rich Facial Expressions," SICE-ICASE Int. Joint Conf. 2006, pp. 5423-5428, 2006.

[3] Business Design Laboratory Co. Ltd., "Communication Robot ifbot," http://www.ifbot.net.

[4] M. Kanoh, S. Kato, and H. Itoh, "Facial Expressions Using Emotional Space in Sensitivity Communication Robot "ifbot," IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 1586-1591, 2004.

[5] M. Kanoh, S. Iwata, S. Kato, and H. Itoh, "Emotive Facial Expressions of Sensitivity Communication Robot "Ifbot"," Kansei Engineering Int., Vol.5, No.3, pp. 35-42, 2005.

[6] M. Gotoh, M. Kanoh, S. Kato, T. Kunitachi, and H. Itoh, "Face Generator for Sensibility Robot based on Emotional Regions," The 36th Int. Symposium on Robotics, 2005.

[7] J. L. Elman, "Finding structure in time," Cognitive Science, Vol.14, pp. 179-211, 1990.

[8] J. L. Elman, E. A. Bates, M. H. Johnson, D. Parisi, and K. Plunkett, "Rethinking Innateness: A Connectionist Perspective on Development," Bradford Books, 1996.

[9] H. Yamada, K. Suzuki, and S. Hashimoto, "Interrelating physical feature of facial expression and its impression," HCS2000-47, 2001.

[10] R. R. Cornelius, "The Science of Emotion; Research and Tradition in The Psychology of Emotion," Prentice Hall College Div., 1995.

[11] P. Ekman, "Unmasking the Face," Prentice-Hall, 1975.

[12] J. P. Guilford. "Psychometric Methods," 2nd edition, McGraw-Hill Inc., 1954.

[13] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro, "A practical experiment with interactive humanoid robots in a human society," Third IEEE Int. Conf. on Humanoid Robots (Humanoids 2003), 2003.

Name:
Yuki Matsui

Affiliation:
Department of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

Address:
Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan

Brief Biographical History:
2008 Received B.E. degree from Nagoya Institute of Technology
2010 Received M.E. degree from Nagoya Institute of Technology

Main Works:
• "Evaluating A Model for Generating Interactive Facial Expressions using Simple Recurrent Network," 2009 IEEE Int. Conf. on Systems, Man, and Cybernetics, pp. 1701-1706, 2009.

Name:
Masayoshi Kanoh

Affiliation:
Assistant Professor, Department of Mechanics and Information Technology, School of Information Science and Technology, Chukyo University

Address:
101 Tokudachi, Kaizu-cho, Toyota 470-0393, Japan

Brief Biographical History:
2004 Received Ph.D. degree from Nagoya Institute of Technology
2004- Assistant Professor, Chukyo University

Main Works:
• "Emotive Facial Expressions of Sensitivity Communication Robot "Ifbot"," Kansei Engineering Int., Vol.5, No.3, pp. 35-42, 2005.

Membership in Academic Societies:
• The Robotics Society of Japan (RSJ)
• The Japanese Society of Public Health (JSPH)
• The Japan Society of Kansei Engineering (JSKE)

Name:
Shohei Kato

Affiliation:
Associate Professor, Department of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

Address:
Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan

Brief Biographical History:
1998- Research Associate, Toyota National College of Technology
1999- Lecturer, Toyota National College of Technology
2002- Assistant Professor, Nagoya Institute of Technology
2003- Associate Professor, Nagoya Institute of Technology

Main Works:
• "Bayesian Method for Detecting Emotion from Voice for Kansei Robots," JSKE J. of Kansei Engineering Int., Vol.8, No.1, pp. 15-22, 2009.

Membership in Academic Societies:
• The Japanese Society for Artificial Intelligence (JSAI)
• The Information Processing Society of Japan (IPSJ)
• The Institute of Electronics, Information and Communication Engineer (IEICE)
• The Robotics Society of Japan (RSJ)
• The Institute of Electrical and Electronics Engineering (IEEE)

Name:
Tsuyoshi Nakamura

Affiliation:
Associate Professor, Department of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

Address:
Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan

Brief Biographical History:

1998 Ph.D., Nagoya Institute of Technology
1998- Research Associate, Nagoya Institute of Technology
2003- Associate Professor, Nagoya Institute of Technology

Main Works:

- "Color transfer based on hierarchical image-region matching with interactive evolutionary computation," Kansei Engineering Int., Vol.5, No.4, pp. 1-10, 2006.

Membership in Academic Societies:

- The Institute of Electrical and Electronics Engineering (IEEE)
- Association for Computing Machinery (ACM)
- The Institute of Electronics, Information and Communication Engineers (IEICE)
- Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT)
- Japan Society of Kansei Engineering (JSKE)

Name:
Hidenori Itoh

Affiliation:
Professor, Department of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

Address:
Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan

Brief Biographical History:

1974- Joined Nippon Telephone and Telegraph Laboratories
1985- Joined The Institute for New Generation Computer Technology
1989- Professor, Nagoya Institute of Technology

Main Works:

- "Rough Set Based FCM Algorithm for Image Segmentation," Int. J. of Computational Science, Vol.1, No.1, pp. 58-68, 2007.