

Computer User Action Learning with Pulse Neural Network and Ultrasonic Wave

Hidetoshi Nonaka and Tsutomu Da-te
Graduate School of Engineering, Hokkaido University,
N13W8, Kita-Ku, Sapporo 060 8628, Japan
+81 11 706 6793, nonaka@main.eng.hokudai.ac.jp

Abstract

In this paper a learning system of user action is proposed. It is constructed by combining an ultrasonic phase measuring method with an artificial pulse neural network by analogy of a physiological model of auditory processing.

Keywords: ultrasound, pulse neural network, phase difference, action learning

1 Introduction

Among various computer input devices, keyboard and mouse have been used for the standard input devices of computer systems. Their performances have been examined and improved as years go on, but several inconvenient characteristics are still remaining. For example, a computer user is required to operate them on a level plane and to sit still in front of the computer, therefore, mouse is sometimes substituted by trackball, pointing-stick, touch-screen, touch-pad, stylus pen, and so on. The purpose of this study is to release computer users from such constraints and to allow them to work in arbitrary styles with free posture. Our system will be also applied to an auxiliary or additional input device of ordinary VDT tasks, the device for users with physical disabilities, and so on.

Ultrasonic waves are used for various methods in position measurement, object identification, object localization, imaging and so on (for example, (marioli 1992), (Figuroa 1991), (Sutherland 1968), (Nonaka 1991, 1993), (Arai 1983), (Wagdy 1987), (Lewis 1998)). These are conceptually classified into two types. One is based on the pressure level of ultrasound, and the other on the time of propagation of ultrasonic wave.

Moreover, the latter is classified into continuous wave type and pulse type. In the case of continuous wave type, the distance is measured by the phase shift between two waves (Nonaka, 1995).

By the way, living organisms recognize sound direction using the difference between two sound signals from a pair of auditory receptors (ears). It is known physiologically that the sound-source localization is achieved by *inter-aural time difference (ITD)* and *inter-aural level difference (ILD)*. For example, in the case of primates, the former is effective at lower frequencies, while the latter is effective at higher frequencies. And in the case of owl, horizontal localization is achieved by the former, while vertical localization is achieved by the latter (Knudsen 1987).

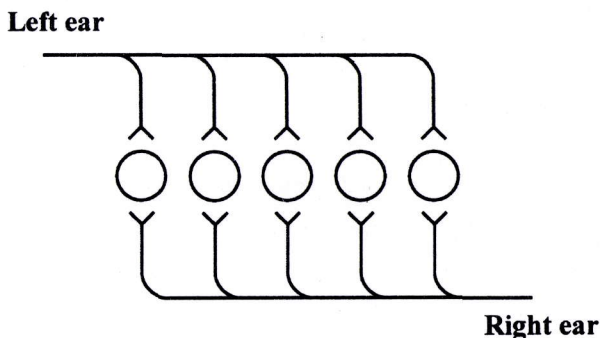


Fig. 1: A model of sound-source localization by inter-aural time difference

This physiological knowledge is applied to sound localization systems, artificial models of the acoustic localization, and so on (Huang 1988), (Kuroyanagi 1994, 1996). In the field of virtual reality, it is utilized to construction of virtual acoustic displays. By translating these descriptions of the audible sound source localization into those of ultrasonic measurement, the techniques of detecting ITD and ILD correspond to those of phase difference and pressure level, respectively.

Yoneyama (1992) proposed an object recognition system by combining imaging by pressure level of ultrasound with neural networks, and in 1997, realized a 3-D object recognition facility with 64 ultrasonic sensors based on a 3-D imaging system (Watanabe 1992). On the other hand, we have proposed a 3-D object identification

system with 3 ultrasonic sensors, combining an ultrasonic phase measuring method with a pulse neural network (Nonaka, 2000).

We have constructed a pulse neural network by hardware implementation, which is based on a simple modeling of the inter-aural phase difference detection, in order to apply directly the information of phase difference to the neural network. Consequently, the number of sensors is reduced by restricting the purpose to only the object identification and omitting the imaging. In the lower part of the network — the first layer and the second layer —, the processing is parallel and asynchronous, therefore, the execution time in the lower part of the network does not increase with the number of sensors and units. But in the second, third, and output layers, the back-propagation method is used with a deterministic feed-forward network. Thus, static objects alone can be identified by the system.

2 Pulse Neural Network

The essential idea of the pulse artificial neural network is the use of architecture and learning paradigms from the biological knowledge. It allows us extracting of not only the spatial structure but also the temporal structure from a set of data without *a priori* knowledge of the object.

Eckhorn, et al. (1990) proposed the pulse coupled neural network (PCNN). The PCNN is a physiologically motivated artificial neural network for visual sensory system that is composed of artificial spiking neurons. It models after the characteristics of the pulse height, duration, repetition rate, and modulatory interneural linking observed in biological dendrites in visual cortex of monkeys and cats. It is utilized in scene segmentation, object detection, image smoothing, feature extraction, and so on (Eckhorn 1999), (Broussard 1999). Interested readers may find various applications of PCNN in the same volume — “Special Issue on PCNN” in *IEEE Trans. on Neural networks* 10(3), 1999 —. Kuroyanagi and Iwata (1994, 1996) proposed an auditory pulse neural network model to extract the Inter-aural time and level difference for sound localization. They constructed models of medial superior olivary nucleus and lateral superior olivary nucleus, which stand for ITD and ILD, respectively. Pulse neural

models are implemented by analog circuits, digital systems, and so on (Frank 1999), (Jahnke 1999), (Ota 1999).

An example of models of pulse neural network is shown in Fig. 2, and the principle of pulse neural network is shown in Fig. 3.

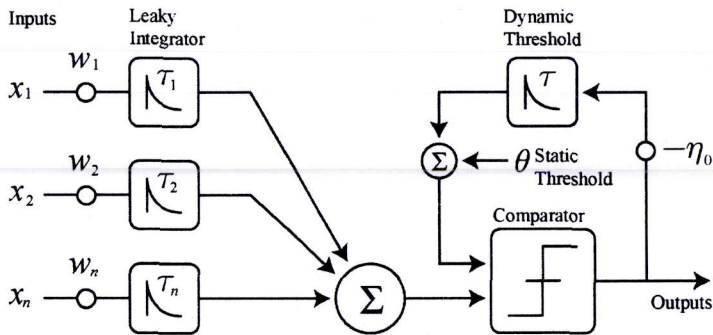


Fig. 2: Example of model neuron.

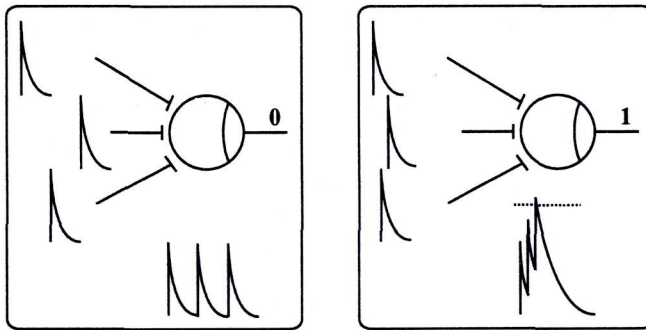


Fig. 3: The principle of pulse neural network.

In this paper we propose an extended system with 8 ultrasonic sensors. Moreover, pulse neural network is applied to not only the lower part but also the higher part of the network, and parameters of each cell are improved by spike-based learning instead of ordinary error back-propagation, in order to identify moving objects and action of computer users.

3 Construction of the System

Eight sensors are arranged equally around a circle with a radius of 18mm, and a transmitter is positioned on the center of the circle. These are embedded in an ultrasound-absorbing sheet. The arrangement is shown in Fig. 4.

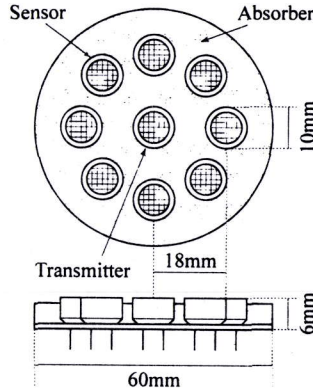


Fig. 4: The arrangement of sensors and a transmitter.

The continuous ultrasonic wave (40KHz) generated by the transmitter reflects on a moving object and reaches the sensors as shown in Fig. 5. The sinusoidal signal of 40KHz from the sensor is converted to pulse-trains $\alpha_i(t)$ by a voltage comparator and function generator in the first layer.

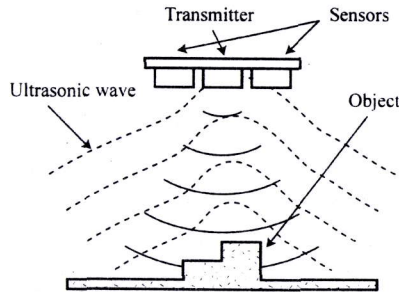


Fig. 5: Transmission and reflection of ultrasonic wave.

The network architecture (first layer and second layer) is shown in Figs. 6-7. Second layer consists of eight groups, and each group consists of 256 units. The pulse trains generated in the first layer are transmitted to the second layer of neural network.

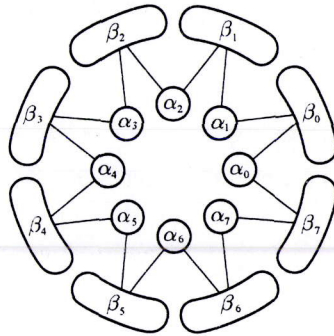


Fig. 6: Connection between 1st layer (α) and 2nd layer (β).

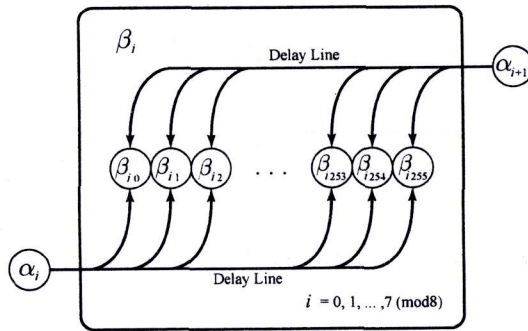


Fig. 7: Connection between 1st layer and 2nd layer (detailed).

The mathematical description is summarized as follows. Though our system simulates the pulse neural network in discrete-time, we stay for a while at the level of the underlying concept with analog description for convenience. When a pulse $\alpha_i(t)$ reaches the j -th unit in i -th group of the second layer from i -th unit in the first layer at t_i , a local potential $\tilde{\beta}_{ij}(t)$ is generated and decays exponentially with time constant τ_1 :

$$\tilde{\beta}_{ij}(t) = e^{-(t-t_i)/\tau_1} \mathcal{H}(t-t_i), \text{ where } \mathcal{H} \text{ represents the Heavyside step function. In the}$$

second layer, local potentials are integrated by $\beta_{ij}(t) = f\left(\sum_{k=i}^{i+1(\text{mod } 8)} w_{kj} \tilde{\beta}_{kj}(t_k)\right)$. w_{kj} is the synaptical strength. $f(\cdot)$ is the activation function: $f(x) = e^{-(t-t_0)/\tau_2}$, where t_0 denotes the time when x exceeds the static threshold θ , which is chosen empirically. τ_2 is the identical time constant of second layer.

These signals from the first layer pass through delay lines. The summation of local potentials increases as the time dispersion decreases, therefore, each unit in the second layer becomes active only when it receives each signal with the specific phase difference.

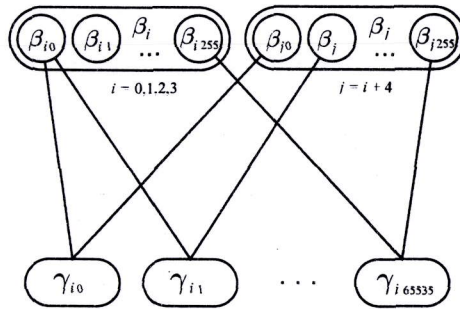


Fig. 8: Connection between 2nd layer (β) and 3rd layer (γ).

In our system, the network is simulated in discrete-time. A continuous time period (40KHz) is divided by a counter (10.24MHz), therefore, the data in the first layer is expressed by an integer from 0 to 255. Each unit in the second layer is connected with two presynaptic units in the first layer, then the time interval between two inputs α_i and α_{i+1} can be used instead of the local potential $\beta_{ij}(t)$, i.e., a spike is generated by

β_{ij} in the second layer when $|\alpha_{i+1} - \alpha_i - d_{ij}|$ becomes smaller than \mathcal{T}_1 , where \mathcal{T}_1 is the time threshold, and d_{ij} is the time delay of the unit β_{ij} .

The network architecture of second and third layers is shown in Figs. 8-9. Third layer consists of four groups, and each group consists of 65536 units. When a pulse $\beta_{i^*}(t)$ reaches the k -th unit in i -th group of the third layer from i -th unit in the second layer at t_i , a local potential $\tilde{\gamma}_{ik}(t)$ is generated and decays exponentially with time constant τ_2 : $\tilde{\gamma}_{ik}(t) = e^{-(t-t_i)/\tau_2}$. In the third layer, local potentials are integrated by $\gamma_{ik}(t) = g\left(\sum_{l=i}^{i+4(\text{mod } 8)} w_{il}\tilde{\gamma}_{il}(t_l)\right)$. $g(\cdot)$ is the activation function: $g(x) = e^{-(t-t_0)/\tau_3}$, where t_0 denotes the time when x exceeds the threshold θ . τ_3 is the identical time constant of the third layer.

Similarly to the first layer, these signals from the second layer pass through delay lines. The summation of local potentials increases as the time dispersion decreases, therefore, each unit in the third layer becomes active only when it receives each signal with the specific phase difference. Moreover, the network is also simulated in discrete-time. Each unit in the third layer is connected with two units in the second layer, therefore, the local potential $\gamma_{ik}(t)$ can be replaced by the time interval of two presynaptic inputs β_i and β_{i+4} .

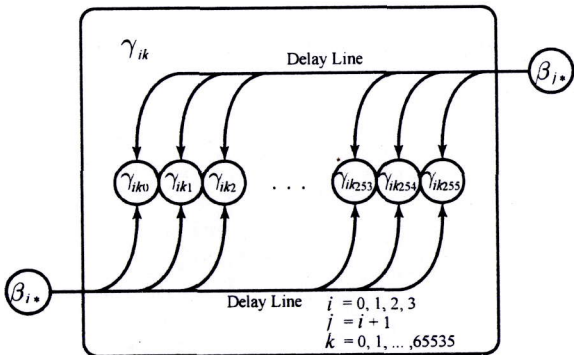


Fig. 9: Connection between 2nd layer and 3rd layer (detailed).

Fig. 10 illustrates the network architecture of third layer and fourth layer. Some of 65536 units are selected from each group of the third layer, and signals from four units are summed up through leaky integrators. Conceptually, in the second layer and the third layer, phase difference is measured by pulse neural network, on the other hand, the frequency of pulses are measured in the fourth layer. The mathematical description is as follows.

$$\delta_x(t) = \sum_{j \in \Gamma_x} \sum_{t_x^{(f)} \in \mathcal{F}} w_{xj} \gamma_{xj}(t - t_x^{(f)})$$

$$\mathcal{F}_x = \{t_x^{(f)}; 1 < f < n\} = \{t \mid \delta_x(t) = \theta\}$$

$$\Gamma_x = \{j \mid j \text{ presynaptic to } x\}$$

where $t_x^{(f)}$: the firing time of neuron x .

\mathcal{F}_x : the set of all firing times of neuron x .

θ : if threshold θ is reached by $\delta_x(t)$, neuron x emits a spike.

Γ_x : the set of neurons j presynaptic to x .

w_{xj} : the synaptical strength.

Each unit δ_x corresponds to a user's action. Conceptually, there are 8, 2048 and 524288 units in the first, second and third layer, respectively. But in our simulation system, 8, $8n$ and $\sum_{x=1}^n \text{card}(\Gamma_x)$ units in each layer are computed for n actions.

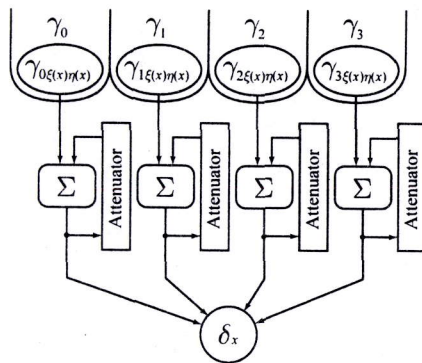












Fig. 10: Connection between 3rd layer (γ) and 4th layer (δ).

4 Experiment

Finger numerals were used in the experiment (Table 1). The 3 sets: "0-9", "0-5" and "6-9" are tested. The right hand was positioned in front of the transmitter at the distance of about 20[cm]. The input data was separated into two sets for each numeral: 100 trials for training and 100 trials for testing. During the experiment, The state of each layer and each unit is monitored through the monitoring window (Figure 11).

Table 1: Finger numerals: 0 - 9

0	1	2	3	4	5	6	7	8	9
									

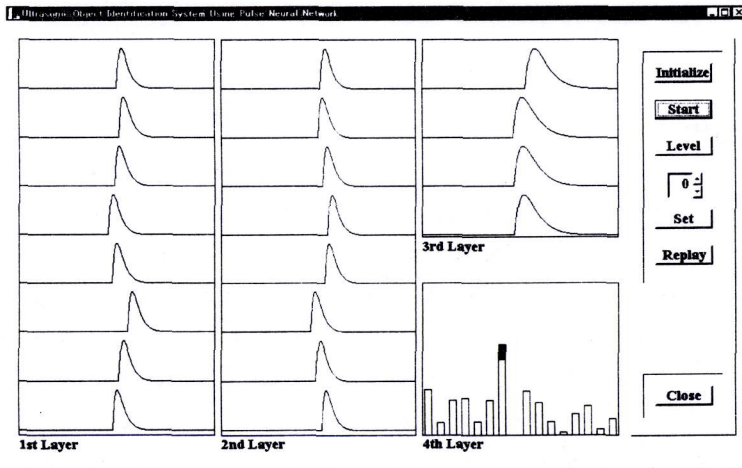


Fig. 11: The monitoring window of pulse neural network

The experimental results are shown in table 2 and 3. Table 1 shows the number of correct identifications for 100 trials for 3 sets. Table 2 shows the duration for identifications.

Table2: Number of correct identifications for 100 trials.

Numerals	0	1	2	3	4	5	6	7	8	9
Set1 "0 - 9"	67	72	68	48	46	67	46	45	29	33
Set2 "0 - 5"	97	95	97	94	92	96	-	-	-	-
Set3 "6 - 9"	-	-	-	-	-	-	52	48	47	40

Table2: Duration for identifications (sec)(average of correct identifications).

Numerals	0	1	2	3	4	5	6	7	8	9
Set1 "0 - 9"	.89	.88	.45	.66	.77	.44	.90	.97	.89	.91
Set2 "0 - 5"	.45	.54	.37	.43	.43	.37	-	-	-	-
Set3 "6 - 9"	-	-	-	-	-	-	.77	.87	.87	.91

5 Conclusion

We proposed a computer user action learning system with pulse neural network and ultrasonic phase difference measurement. It is composed of eight ultrasonic sensors and four-layered pulse neural network. The network is simulated in discrete time. In an experiment, finger language — a set of manual numerals — is tested in an environment of virtual VDT tasks. The experimental results show that the proposed method is profitable for menu-driven graphical user interface (up to 5 menus). Further research is in progress on applying the system to more complex tasks, for example, text editing, and evaluation of the performance in real computer operations. Moreover, some extensions are under way: 1) refractory period to each unit; 2) Hebbian learning to the synaptical strength; and 3) lateral inhibition to the network structure.

References

1. Arai, T., and Nakano, E. (1983): Development of Measuring Equipment for Location and Direction (MELODI) Using Ultrasonic Waves, *Trans. of the ASME, J. DSMC*, Vol. **105**, pp. 152-156.
2. Broussand, R. P., Rogers, S. K., Oxley, M. E., and Tarr, G. L. (1999): Physiologically Motivated image Fusion for Object Detection using a Pulse Couple Neural Network, *IEEE Trans. on Neural Networks*, Vol. **10**, No. 3, pp. 554-563.
3. Eckhorn, R., Reitboeck, H. J., Arndt, M., and Dicke, P. (1990): Reature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex, *Neural Computation*, Vol. **2**, pp.293-307.
4. Eckhorn, R. (1999): Neural Mechanisms of Scene Segmentation: Recordings from the Visual Cortex Suggest Basic Circuits for Linking Filed Models, *IEEE Trans. on Neural Networks*, Vol. **10**, No. 3, pp. 464-479.
5. Figueroa, F., and Barbieri, E. (1991): An Ultrasonic Ranging System for Structural Vibration Measurements, *IEEE Trans. on Instrumentation and Measurement*, Vol. **40**, No. 4, pp. 764-769.
6. Frank, G., Hartmann, G, Jahnke, A., and Schäfer, M. (1999): An Accelerator for Neural Networks with Pulse-Coded Model Neurons, *IEEE Trans. on Neural Networks*, Vol. **10**, No. 3, pp. 527-538.
7. Huang, J, Ohnishi, N., and Sugie, N. (1988): Research on a Sound Source Localization System Suggested by Living Organism — Localization of a Single Sound Source in an Anechoic Environment —, *IEICE Trans. A*, Vol. **J71-A**, pp. 1780-1789 (in Japanese).
8. Jahnke, A., Roth, U., and Schönauer, T. (1999): Digital Simulation of Spiking Neural Networks, in "Pulsed Neural Networks", Maass, W., and Bishop, C., M., Eds., MIT Press.
9. Knudsen, E. I. (1987): The Hearing of the Barn Owl, *Scientific American*, Vol. **245**, pp.83-91.
10. Kuroyanagi, S., and Iwata, A. (1994): Auditory Pulse Neural Network Model to Extract the Inter-Aural Time and Level Difference for Sound Localization, *IEICE Trans. INF. & SYST.*, Vol. **E77-D**, No. 4.
11. Kuroyanagi, S., and Iwata, A. (1996): Auditory Pulse Neural Network Model for Sound Localization — Mapping of the ITD and ILD —, *IEICE Trans. D-II*, Vol. **J79-D-II**, pp. 267-276 (in Japanese).

12. Lewis, J. T., Galloway, R. L., and Schreiner, S. (1998): An Ultrasonic Approach to Localization of Fiducial Markers for Interactive, Image-Guided Neurosurgery, *IEEE Trans. on Biomedical Engineering*, Vol. **45**, No. 3, pp. 620-630.
13. Maass, W., and Bishop, C. M., Eds. (1999): Pulsed Neural Networks, MIT Press.
14. Marioli, D., Narduzzi, C., Offelli, C., Petri, D., Sardini, E., and Taroni, A. (1992): Digital Time-of-Flight Measurement for Ultrasonic Sensors, *IEEE Trans. on Instrumentation and Measurement*, Vol. **41**, No. 1, pp. 93-97.
15. Nonaka, H. and Da-te, T. (1995): Ultrasonic Position Measurement and its Applications to Human Interface, *IEEE Trans. on Instrumentation and Measurement*, Vol. **44**, No. 3, pp.771-774.
16. Nonaka, H. and Da-te, T. (1993): Development of a System for Real Time Position Measuring of Viewpoint, *Human Interface News and Report*, Vol. **8**, pp. 177-182 (in Japanese).
17. Nonaka, H. and Da-te, T. (1991): Video System with Properties of Binocular and Motor Vision, *Trans. of the Society of Instrument and Control Engineers*, Vol. **27**, No.10, pp.1144-1151 (in Japanese).
18. Nonaka, H., Nakashima, K, and Da-te T. (2000): An Ultrasonic 3-D Object Identification System Using Pulse Neural Network, *AIP Conference Proceedings: CASYS'99*, Vol. **517**, pp.621-627.
19. Ota, Y, and Wilamowski, B. M (1999): Analog Implementation of Pulse-Coupled Neural Networks, *IEEE Trans. on Neural Networks*, Vol. **10**, No. 3, pp. 539-544.
20. Sutherland, I. (1968): A Head-Mounted Three Dimensional Display, *Fall Joint Computer Conference*, pp. 757-764.
21. Wagdy, M. F., and Lucas, M. S. P. (1987): A Phase-Measurement Offset Compensation Technique Suitable for Automation, *IEEE Trans. Instrumentation and Measurement*, Vol. **IM-36**, No. 3, pp. 721-724.
22. Watanabe, S., and Yoneyama, M. (1992): An ultrasonic visual sensor for three-dimensional object recognition using neural networks, *IEEE Trans. on Robotics and Automation*, Vol. **8**, No. 4, pp. 40-49.
23. Yoneyama, M., Yuasa, K., and Horie, J. (1997): Post Processing for 3-D Ultrasonic Imaging System Using Associative Memory Effect of Hopfield Neural Networks, *IEICE Trans. D-II*, Vol. **J80-D-II**, pp. 2958-2968 (in Japanese).