

Paper:

Nonverbal Communication Based on Instructed Learning for Socially Embedded Robot Partners

Ryosuke Tanaka, Jinseok Woo, and Naoyuki Kubota

Intelligent Mechanical Systems, Graduate School of System Design, Tokyo Metropolitan University

6-6 Asahigaoka, Hino, Tokyo 191-0065, Japan

E-mail: {tanaka-ryosuke@ed., woojs@, kubota@}tmu.ac.jp

[Received November 30, 2018; accepted February 19, 2019]

The research and development of robot partners have been actively conducted to support human daily life. Human-robot interaction is one of the important research field, in which verbal and nonverbal communication are essential elements for improving the interactions between humans and robots. Thus, the purpose of this research was to establish a method to adapt a human-robot interaction mechanism for robot partners to various situations. In the proposed system, the robot needs to analyze the gestures of humans to interact with them. Humans have the ability to interact according to dynamically changing environmental conditions. Therefore, when robots interact with a human, it is necessary for robots to interact appropriately by correctly judging the situation according to human gestures to carry out natural human-robot interaction. In this paper, we propose a constructive methodology on a system that enables nonverbal communication elements for human-robot interaction. The proposed method was validated through a series of experiments.

Keywords: robot partner, human-robot interaction, cognitive development system, instructed learning, information support

1. Introduction

Research on cognitive science based on physicality has recently began to flourish again for the first time since the 1980s. Cognitive development in areas such as communication skills has been researched from various viewpoints [1, 2]. Communication methods can be classified as verbal and nonverbal communication. When communicating with communication partners, communication methods using verbal information are efficient in most cases. However, nonverbal communication plays an important role in collaborative team working [3–5]. In order to systematically understand communication, communication methods are being researched from various viewpoints in fields such as philosophy and pragmatics. Research on nonverbal communication has provided knowledge useful for studies on human-robot interactions [6].

For example, Michael Tomasello's idea about cognitive development was influenced by Grice's cooperative principle and Sperber and Wilson's relevance theory [7, 8].

Human communication is fundamentally a cooperative act. It is believed that, during the course of human evolution, individuals who could communicate with each other in a collaborative, cooperative, and motivational way were at the forefront of adaptation. Such theories naturally provide important knowledge regarding communication design for human-robot interactions. In addition, it is believed that verbal communication began to be used after nonverbal communication [9]. Moreover, studies on cognitive developmental robotics have proposed various methods of cognitive developmental learning based on physicality based communication [10]. Development in the form of the mutual involvement of cognitive abilities, communication abilities, and physical exercise ability is considered to be important for human growth.

As described above, it is important to simultaneously learn verbal information and nonverbal information in the process of acquiring language during human cognitive development. Behavior learning during cognitive development can be broadly classified into social learning, which is mediated by others, and self-learning by trial and error. Furthermore, in social learning, unidirectional learning by imitating others and bidirectional learning based on teaching can be considered.

In this study, we aim to clarify the interdependent time-dependent learning processes of physical exercise ability and communication ability in cognitive development through interactive learning based on human teaching between humans and humanoid robots. The rest of this paper is organized as follows. Section 2 describes the robot partner system; Section 3 explains the proposed instructed learning system for the robot partners; Section 4 describes our methodology for dealing with human motion based on a spiking neural network (SNN), a self-organizing map (SOM), and a growing neural gas (GNG). Experimental results with robot partners are shown in Section 6. Finally, Section 7 highlights some of the future directions of our research.



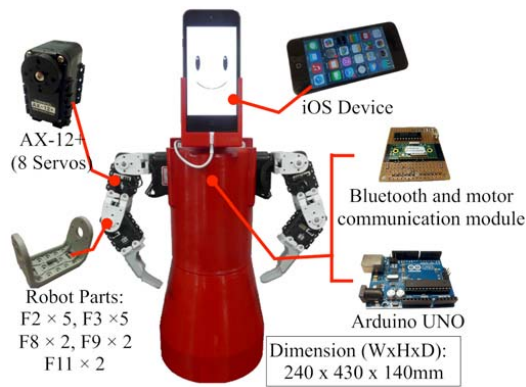


Fig. 1. Robot partner: iPhonoid-C.

2. Robot Partner System

Robots are going to become familiar in our daily lives [11], especially in the service industry. Robots may be one of the solutions to modern social problems due to the shortage of labor. In general, robots in the current market face various restrictions. Robots have not been made widely available because of their limitations, such as their high initial startup cost in term of hardware design and software development. In order to overcome these issues, we propose a robot partner using a modular structure (Fig. 1 [12, 13]). The proposed robot partner consists of a smartphone, a body structure that can be fabricated using a 3D printer, a microcontroller, and several servo motors [14]. Each module can be designed and reproduced on the spot. For example, users can create their own robot partner by installing the proposed application into their smart devices, then select the desired 3D design for the robot's body, and finally send the selected design to a 3D printer for printing. The following sections describe the architecture of the robot partner design.

2.1. Architecture of the Robot Partner

Robot partner prototypes have been developed for different objectives and considering various improvements [12]. Therefore, our robot partner was designed to have a modular hardware structure as shown in Fig. 2. The hardware of the robot can be printed by a 3D printer with various degree of freedom (DOF) for gesture generation as shown in Fig. 1 (e.g., neck, arms, and waist). These robot partners are equipped with several servo motors. Using this system structure, robot gestural generation can be realized. In addition, an Arduino microcontroller board with Bluetooth capabilities is attached to the body of the robot partner for controlling its servo motors through an iOS device [15, 16]. The robot partner system has a modular structure to allow for a modular cognitive model [17]. Each module consists of four main components for developing the essential functions of the robot: hardware, low-level software library, electronic components, and high-level software library (iOS). These four main components can be used to develop a

	Hardware Layer	Library Layer	Component Layer	Application Layer
iOS Device Part	Display	OpenCV	Voice recognition	Human-robot interaction
	Touch sensor	OpenGL ES	Speech synthesis	Gesture generation
	Mic	MySQL Client	Natural language processing	Multi language
	Speaker	UIKit	State estimation	Exercise assistance
	Camera	Foundation	Emotional model	Information support
	Proximity sensor	AVFoundation	Laban movement	Elderly care system
	Motion sensor/ accelerometer	CoreBluetooth	Facial expression	Hotel reception
	Gyroscope	AudioToolbox	Human detection	Emergency support
	Ambient light sensor	CoreImage	Face detection	Recreation
	Compass	CoreMedia	Gesture recognition	Schedule reminder
	Battery state	Security	Object tracking	Robot theater
		ExternalAccessory	Gender classifier	
		Speech		
Robot Body Part	3D Printing model	Actuator control	Movement control	Countenance
	Actuators	Serial communication	Motor state	Robot behavior
	Micro controller		Bluetooth communication	Gesture expression
	Bluetooth LE			

Fig. 2. Modular structure of the proposed robot partner system.

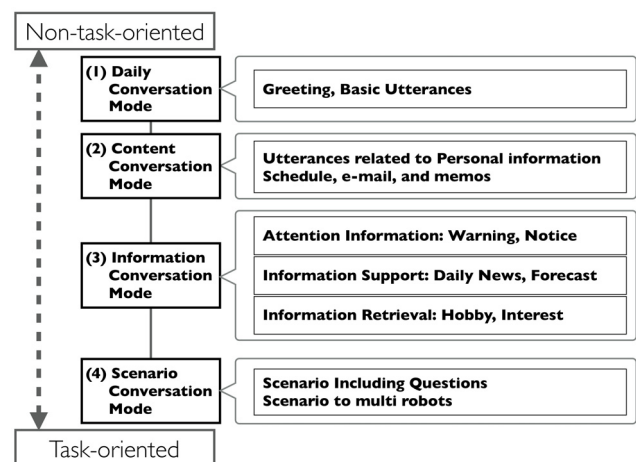


Fig. 3. Structure of the time dependent utterance system.

robot partner capable of performing tasks dictated by humans.

2.2. Conversation System

In this section, we explain the conversation modes of the proposed robot partner. We developed a system that focuses on time-dependent utterances for human-robot conversation [18, 19]. The conversation modes can be classified into four types, namely daily conversation mode, contents conversation mode, information conversation mode, and scenario conversation mode as shown in Fig. 3.

In daily conversation mode, a conversation is carried out based on the user's daily lifestyle, which is extracted from their daily lifelog. In content conversation mode, a conversation is carried out using information from the smartphone, such as the calendar and e-mail. In addition, these conversations can serve as reminders according to the user's schedule. This mode partly overlaps with the information conversation mode. In this mode, a conversation is carried out based on news, the weather forecast,

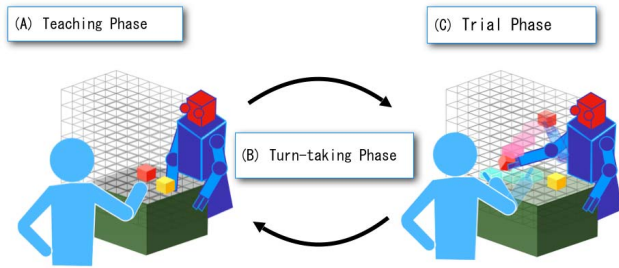


Fig. 4. Instructed learning.

schedule management, and hobby information regarding interesting topics for the elderly [20]. Lastly, in scenario conversation mode, the system provides varied information in a dialogue form for the elderly. In this mode, conversation packages are prepared according to certain purposes, such as “hotel receptionist support” or “elderly care support.” In this paper, we analyze nonverbal communication in the form of human gestures based on the above-described robot conversation system.

3. Instructed Learning for Nonverbal Communication

Instructed learning is conducted for robot partners with different levels of knowledge and skills [21]. Through instructed learning in human-robot interactions, the system performs the symbolization of internal perceptions and action information based on nonverbal communication in a mutual cognitive environment. This study was also aimed at elucidating the learning process of languageization to prompt explicit externalization from a constitutive point of view. Teaching motions considered include object-motion operations, such as those performed with building blocks and balls of different shapes and colors.

In order to develop an instructed learning system, we divide the teaching phase into three phases, namely the teaching phase for recreating target actions performed by human teachers; the trial phase, in which the learner performs a trial; and the turn-taking phase to switch between these two phases, as shown in Fig. 4. Furthermore, for each of these three phases, a clarification of the necessary perceptual and action functions is presented in Table 1. The learner’s (robot) instructed learning system is formulated in consideration of the necessary functional configuration described above. We developed a method to recognize a target action, which is the functional configuration required for the learner and extract the target action segment.

4. Recognition of Target Motion

In this section, we describe a target motion analysis method using an SNN for robot motion recognition.

4.1. Spiking Neural Network for the Recognition of Motion Patterns

We employed an SNN for recognizing human motion patterns [22]. An SNN is a type of neural network model that focuses on internal potential neurons [23]. It mimics the mechanisms of pulse propagation between nerve cells in the human brain. SNNs artificially reproduce the action potentials emitted by nerve cells, thereby creating an input and output of temporal pulses on the neural network. Because SNNs have a higher fidelity for biomimetic simulations and better time series recognition abilities than conventional neural networks, they are now widely used for tasks such as speech and gesture recognition for robots [24].

In SNNs, the internal state (membrane potential) $h_i(t)$ of the i -th neuron at discrete time t is described by the following equation:

$$h_i(t) = \tanh(h_i^{syn}(t) + h_i^{ext}(t) + h_i^{ref}(t)), \quad \dots \quad (1)$$

$h_i^{syn}(t)$ is calculated as follows:

$$h_i^{syn}(t) = \gamma^{syn} \cdot h_i(t-1) + \sum_{j=1}^N w_{ji} \cdot p_j(t-1), \quad \dots \quad (2)$$

where γ^{syn} is the temporal discount rate of h_i and w_{ji} is the weight from the j -th neuron to the i -th neuron. To prevent bursts of neuronal firings, the hyperbolic tangent function was used, as shown in Eq. (1). When the internal state of the i -th neuron reaches a predefined threshold level, a pulse is outputted as follows:

$$p_i(t) = \begin{cases} 1 & \text{if } h_i(t) \geq \theta, \\ 0 & \text{otherwise,} \end{cases} \quad \dots \quad (3)$$

where θ is the threshold for firing. In case of firing, R is subtracted from the $h_i^{ref}(t)$ value of neuron i :

$$h_i^{ref}(t) = \begin{cases} \gamma^{ref} \cdot h_i^{ref}(t-1) - R & \text{if } p_i(t-1) = 1, \\ \gamma^{ref} \cdot h_i^{ref}(t-1) & \text{otherwise,} \end{cases} \quad \dots \quad (4)$$

where $R > 0$, γ^{ref} is a discount rate of h_i^{ref} , and $0 \leq \gamma^{ref} \leq 1$. The coupling strength weight between neurons, w_{ji} , is updated based on learning using Hebb’s rule.

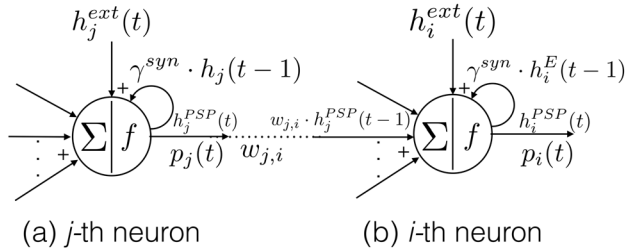
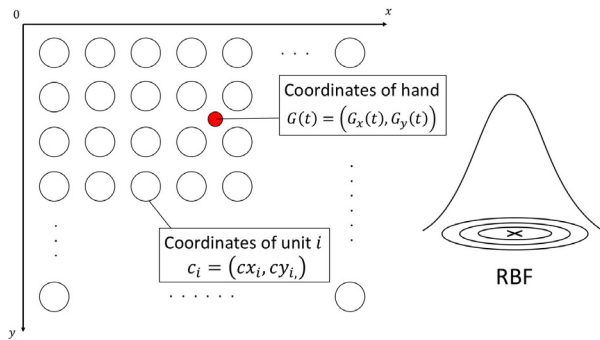
$$w_{ji} \leftarrow \gamma^{syn} \cdot w_{ji} + \xi^{wgt} \cdot p_i(t) \cdot p_j(t-1) \quad \dots \quad (5)$$

$h_i^{ext}(t)$ represents external input. By setting $h_i^{ext}(t)$, it is possible to learn using a SNN designed for specific purposes.

The postsynaptic potential is the response of a postsynaptic neuron to a presynaptic action potential [23]. This is illustrated in Fig. 5, where the j -th neuron is the presynaptic neuron (sender) and the i -th neuron is the postsynaptic neuron (receiver). The presynaptic spike output is transmitted to the connected neuron through a weighted connection (Fig. 5).

Table 1. Required functional configuration of each phase.

Phase	Instructor	Learner
Instruction phase	Sharing cognitive environment Sharing learning target action	Recognition of target motion Extract target action segmentation Virtual reproduction of target motion
Turn-taking phase	Handover of target object Distribution of gaze	Nonverbal information extraction Extract start / end information
Trial phase	Virtual replay of target action Confirm target action	Reproduction by imitation Instructor information extraction

**Fig. 5.** The structure of neurons in a spiking neural network.**Fig. 6.** Neuronal arrangement of the proposed SNN and external input to neuron.

4.2. Motion Recognition Method

In the proposed system, human motion is detected using the coordinates of the hand on the 2D plane, which are obtained via a Kinect sensor [25] as the external input for the SNN. As shown in **Fig. 6**, neurons are arranged in a two-dimensional plane. The input $h_i^{ext}(t)$ from the environment to the i -th neuron at a certain discrete time t is derived using the radial basis function (RBF).

$$h_i^{ext}(t) = \exp\left(-\frac{\|c_i - G(t)\|^2}{\sigma^2}\right) \cdot \dots \cdot (6)$$

$c_i = (cx_i, cy_i)$ are the coordinates of the i -th neuron on the image and σ is the standard deviation. The time series of the pulse output $p_i(t)$ is obtained from the coordinates of the human hand: $G(t) = (G_x(t), G_y(t))$, $(t = 1, 2, \dots, T)$.

5. Target Action Segmentation

There are both meaningless and meaningful movements in the actions that humans perform in their daily lives. Furthermore, meaningful actions are often understood differently depending on the situation.

As shown in Section 4.1, it is possible to recognize motions based on the neuronal firings of an SNN, which match the trajectory of the hand position. However, SNNs recognized not only meaningful actions, but also meaningless ones. Therefore, even if a neuron actually fires, it is difficult to judge whether the action is a meaningful action or not. Consequently, after recognizing the human action, the robot should judge the boundaries that determine the range of meaningful actions. Separating a series of human actions and dividing them into multiple actions is termed as “action segmentation.”

In the following sections, we present a brief introduction to SOMs and GNGs. Then, we describe action segmentation based on these approaches.

5.1. Self Organizing Maps (SOM)

In order to perform data analysis, it is necessary to extract valuable information from the data. In order to use a suitable clustering method for our research, it is necessary to know the characteristics of the acquired data. However, it is generally difficult to visualize multidimensional data, especially those with four or more dimensions. Consequently, SOMs are one of the most frequently used approaches for such data [26]. SOMs represent an unsupervised learning technique that can perform topological mapping in a lower dimensional space by using a reference vector on a multidimensional space. In addition, SOMs have the advantage that they can easily be applied to various kinds of problems by determining the number of nodes and the neighborhood structure.

When a SOM is given an input vector, the node that is closest to the input vector in terms of Euclidean distance is selected as the winning node. Then, learning is performed by updating the reference vector w so that the winner node and its surrounding nodes approach the input vector v . The update process for the reference vector w is as follows:

$$w \leftarrow w + \exp\left(-\frac{dis^2}{(1-\frac{t}{T})}\right) \cdot (v - w) \cdot \dots \cdot (7)$$

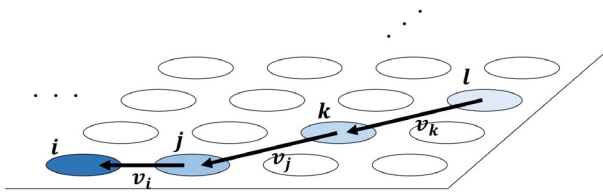


Fig. 7. Calculation of the direction vector.

dis represents the distance between the reference vector and the winner node, t represents the current number of learning instances, and T represents the previous number of learning instances. This process approaches the input vector v according to the distance between two nodes and the current number of learning instances.

5.2. Growing Neural Gas (GNG)

SOMs offer the advantage that they can relatively easily be applied to various kinds of problems by determining the number of nodes and the neighborhood structure. However, it should be noted that it is difficult to select an appropriate number of nodes and the neighborhood structures for unknown data. Moreover, although there are representative methods of unsupervised classification, such as the K -means algorithms [27], it is necessary to select the number of clusters beforehand. GNGs represent an unsupervised classification method that solves these problems [28]. A GNG starts learning using a minimum number of nodes, and then adds/deletes connections between its nodes according to the structure of the input data and performs unsupervised classification. Therefore, a GNG is a self-organizing incremental neural network that can operate using unknown data.

In this study, we used a multi-layer GNG (MLGNG) that can perform hierarchical clustering. In the proposed model, a GNG at a higher level repeatedly clusters the output from a GNG at a lower level. This makes it possible to perform clustering for unknown data with different the phase composition granularities.

5.3. Action Segmentation Method

Human actions are detected using an SNN, as mentioned in Sections 4.1 and 4.2. Using Hebb's learning rules, the SNN can memorize the temporal firing patterns of various actions via neurons associated to the trajectory of the observed human hands. The firing patterns of the SNN are shown in **Fig. 7**. The coordinates of the neuron with the maximum internal potential are used to calculate the direction of vector v_i from the i -th neuron at time t to the j -th neuron at time $t - 1$. Furthermore, the transition probability p_{xi} from the other direction vector v_x to the direction vector v_i is computed.

It is assumed that the obtained direction vector consists of one operation for three units of time, and thus the input vector V to the SOM can be determined.

$$V = (v_i, v_j, v_k) \dots \dots \dots (8)$$

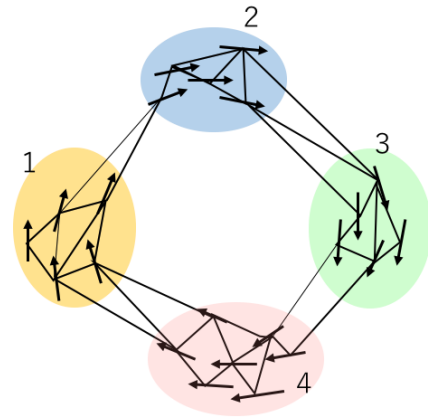


Fig. 8. Clustering using GNG.

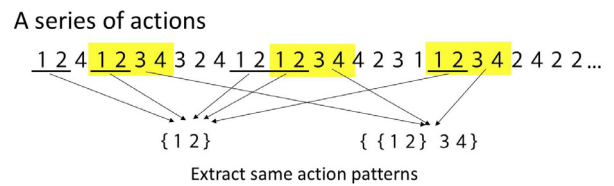


Fig. 9. Extraction of an action pattern via the transition of direction vectors.

The clustering of direction vectors via GNG is performed using the calculated direction vectors as inputs (**Fig. 8**). Similar directional vectors can be grouped as one cluster, and continuous motion can be represented by the transition of direction vectors (**Fig. 9**). In fact, there are no physical breaks between consecutive actions, and thus the system must find meaningful actions in consecutive actions. The cues for finding meaningful actions are the transition probabilities and frequent patterns between actions. Therefore, by extracting the same action pattern using the clustering results of the direction vector obtained via the GNG, actions can be obtained.

6. Experiments

An experiment was conducted using a robot partner and several volunteering candidates for evaluating the performance of the proposed learning system through nonverbal communication. The participants in the experiments were asked to perform several specified actions in a way in which they could interact with the robot partner for us to determine if the system was able to recognize the performed actions. Because the recognition of the SNN was performed on a two-dimensional plane obtained from a camera image, we instructed the participant to operate while taking this into consideration consciousness as much as possible.

For this experiment, we used a Kinect sensor to acquire the trajectory of fingers and feed the data into the learning system to assess the proposed nonverbal com-

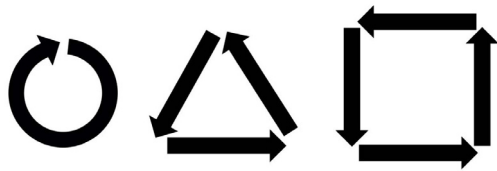


Fig. 10. Action patterns.

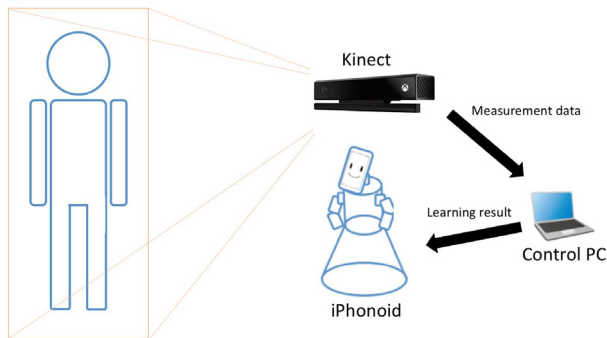


Fig. 11. Overall configuration of the system.

munication learning method. Using the coordinates of the imaged hand as input, motion recognition was performed with the aforementioned SNN. Clustering was performed via the GNG using the direction vector obtained from the recognition results of the SNN. The number of neurons in the SNN was 100, and the setup was placed on a two-dimensional plane with a size of 10×10 .

The hand actions performed by the participants consisted in drawing three different patterns, namely a circle, a square, and a triangle, as shown in Fig. 10. The learning system would then classify the hand motion using the clustering results of the GNG.

A conceptual diagram of the proposed system is illustrated in Fig. 11. The data obtained with the Kinect sensor is used by the robot, which responds according to the learning results obtained from the control PC. The robot is used as a target for communication and the Kinect sensor is deployed behind the robot. Therefore, simulating interaction with robots, human can make gestures. Fig. 12 shows the setup of the experiment.

The motions were performed by the participants over the course of two minutes; these motions consisted of four times circle, two times square, and two times triangle. The neuron firing patterns for each type of motion are shown in Fig. 13. Fig. 13 also shows the results for action recognition of the SNN using the hand position coordinates obtained via the Kinect sensor.

As shown in Fig. 13, it can be confirmed that neuron firings occur along the trajectory of the hand, and in Table 2, the recognition rate achieved was 75%. This shows that the system it is able to recognize multiple action using the SNN. Fig. 14 shows the results of clustering obtained via the GNG, with the direction vector calculated using the internal potential of the SNN as an input.



Fig. 12. Example of an experimental setup.

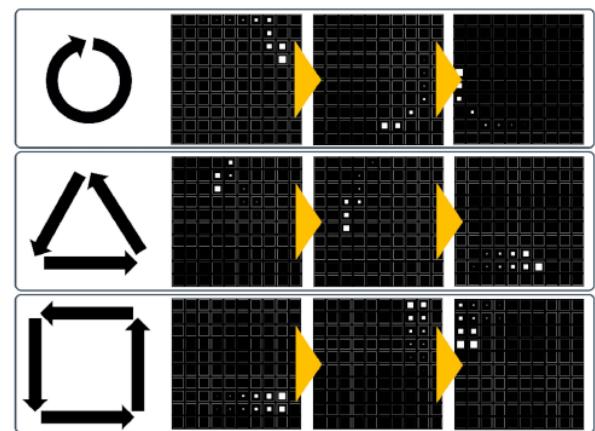


Fig. 13. Action recognition results obtained via the SNN.

Table 2. Required functional configuration of each phase.

Move Pattern	number	correct recognition
Circle	4	3
Square	2	1
Triangle	2	2

In Fig. 14, nodes are represented as direction vectors in a circle, and the connections in the GNG are represented by black lines. The human action pattern is extracted accordingly from the clustering results of the GNG. Series of actions, including the taught motions performed by humans are represented by transitions of the direction vector. The results after extracting the transition patterns from the direction vectors as shown in Fig. 15, representing circles, triangles, and squares.

From Fig. 15, circular, triangular, and square trajectories can be seen. These represent the transitions of the direction vector obtained via clustering using the GNG.

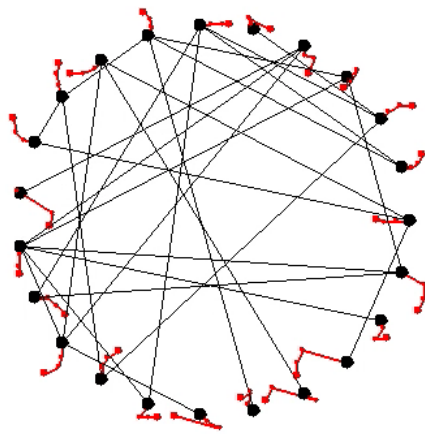


Fig. 14. Clustering results for the direction vector obtained via the GNG.

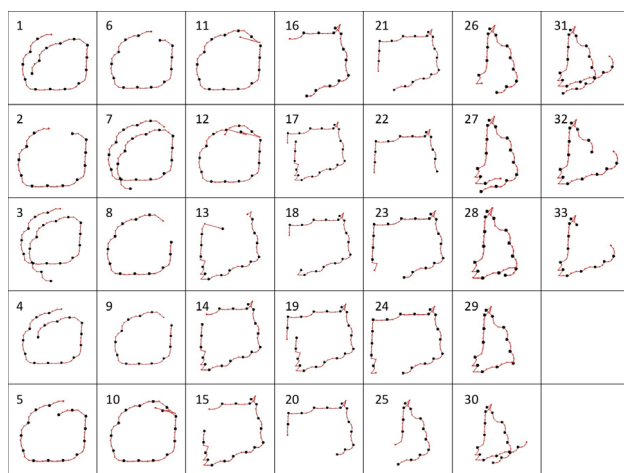


Fig. 15. Clustering results for the action patterns.

7. Conclusions

In this paper, we proposed an instructed learning method for target actions using a nonverbal communication system to realize language acquisition, enabled by instructed learning and target action segmentation. Using a GNG, we can estimate the trajectory of hand motion in a finite number of directions.

We showed that an action can be obtained from the transition patterns and frequent patterns of the direction vector by expressing series of actions as a sequence of direction vectors using the classification results. Because we performed experiments with neurons placed in a 10×10 grid, we consider that it will be possible to obtain more precise results by increasing the number of neurons in the arrangement, thus raising the resolution for recognition. In future works, we will first improve the proposed system for actual interactions with a robot partner. In **Fig. 15**, there are multiple action patterns shown that are similar to the human actions detected using the proposed method. For example, we will extend the system to re-

ject preserved acquired actions through interactions with the robot. Secondly, we will also consider building other learning structures, such as forgetting, to manage other similar operations via human gestures. Thirdly, we will increase the number of motion patterns for learning and recognition and assess the acquisition of actions when the number of clusters of direction vectors increases. Moreover, we will implement a labeling system for acquiring the meaning of words. Then, we will establish a method for attaching language labels to the gestures obtained via the proposed method. Finally, we will extend the system to recognize spatial action by expanding the arrangement of neurons to three-dimensional space.

Acknowledgements

This work was supported by JSPS KAKENHI (Grant-in-Aid for Challenging Exploratory Research) Grant Number 16K12495 and [Grant-in-Aid for Scientific Research on Innovative Areas (Research in a proposed research area)] Grant Number 18H05079.

References:

- [1] K. W. Fischer, "A theory of cognitive development: The control and construction of hierarchies of skills," *Psychological Review*, Vol.87, No.6, p. 477, 1980.
- [2] M. H. Bickhard and L. Terveen, "Foundational issues in artificial intelligence and cognitive science: Impasse and solution," Vol.109, Elsevier, 1996.
- [3] C. Breazeal, C. D. Kidd, A. L. Thomaz, G. Hoffman, and M. Berlin, "Effects of nonverbal communication on efficiency and robustness in human-robot teamwork," 2005 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS 2005), pp. 708-713, 2005.
- [4] A. G. Brooks and R. C. Arkin, "Behavioral overlays for non-verbal communication expression on a humanoid robot," *Autonomous Robots*, Vol.22, No.1, pp. 55-74, 2007.
- [5] A. Mehrabian, "Nonverbal communication," Routledge, 2017.
- [6] C. Rich, B. Ponsler, A. Holroyd, and C. L. Sidner, "Recognizing engagement in human-robot interaction," 2010 5th ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI), pp. 375-382, 2010.
- [7] M. Tomasello, M. Carpenter, and U. Liszkowski, "A new look at infant pointing," *Child Development*, Vol.78, No.3, pp. 705-722, 2007.
- [8] B. L. Davies, "Grice's cooperative principle: meaning and rationality," *J. of Pragmatics*, Vol.39, No.12, pp. 2308-2331, 2007.
- [9] M. Tomasello, "Origins of human communication," MIT Press, 2010.
- [10] R. W. Gibbs Jr., "Embodiment and cognitive science," Cambridge University Press, 2005.
- [11] M. Hirano, N. Hanajima, K. Urita, S. Muto, Y. Muraoka, and M. Ohata, "Development of an Exercise Support System for the Elderly Which Uses a Small Humanoid Robot," *J. Robot. Mechatron.*, Vol.25, No.6, pp. 939-948, 2013.
- [12] J. Woo, J. Botzheim, and N. Kubota, "System Integration for Cognitive Model of a Robot Partner," *Intelligent Automation & Soft Computing*, pp. 1-14, DOI: 10.1080/10798587.2017.1364919, 2017.
- [13] J. Woo and N. Kubota, "Interaction content design for information support based on robot partner," 2017 10th Int. Conf. on Human System Interactions (HSI), pp. 155-160, 2017.
- [14] J. Woo, J. Botzheim, and N. Kubota, "Facial and gestural expression generation for robot partners," 2014 Int. Symp. on Micro-NanoMechatronics and Human Science (MHS), pp. 1-6, 2014.
- [15] S. A. Arduino, "Arduino," Arduino LLC, 2015.
- [16] Apple Inc., "Developer (iOS system)," 2018.
- [17] J. Woo, J. Botzheim, and N. Kubota, "A Modular Cognitive Model of Socially Embedded Robot Partners for Information Support," *Robomech J.*, Vol.4, No.1, p. 10, 2017.
- [18] J. Woo, J. Botzheim, and N. Kubota, "Verbal conversation system for a socially embedded robot partner using emotional model," 2015 24th IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN), pp. 37-42, 2015.
- [19] J. Woo, J. Botzheim, and N. Kubota, "A Socially Interactive Robot Partner Using Content-Based Conversation System for Information Support," *J. Adv. Comput. Intell. Intell. Inform.*, Vol.22, No.6, pp. 989-997, 2018.

- [20] J. Szeles, N. Kubota, and J. Woo, "Weather forecast support system implemented into robot partner for supporting elderly people using fuzzy logic," 2017 Joint 17th World Congress of Int. Fuzzy Systems Association and 9th Int. Conf. on Soft Computing and Intelligent Systems (IFSA-SCIS), pp. 1-5, 2017.
- [21] P. Robinson, "Individual differences and instructed language learning," Vol.2, John Benjamins Publishing, 2002.
- [22] N. Kubota, Y. Tomioka, and M. Abe, "Temporal coding in spiking neural network for gestures recognition of a partner robot," SCIS&ISIS 2006, pp. 737-742, 2006.
- [23] W. Gerstner and W. M. Kistler, "Spiking neuron models: Single neurons, populations, plasticity," Cambridge University Press, 2002.
- [24] S. Loisel, J. Rouat, D. Pressnitzer, and S. Thorpe, "Exploration of rank order coding with spiking neural networks for speech recognition," 2005 IEEE Int. Joint Conf. on Neural Networks (IJCNN'05), Proc., Vol. 4, pp. 2076-2080, 2005.
- [25] Z. Zhang, "Microsoft kinect sensor and its effect," IEEE Multimedia, Vol.19, No.2, pp. 4-10, 2012.
- [26] T. Kohonen, "Exploration of very large databases by self-organizing maps," Proc. of Int. Conf. on Neural Networks (ICNN'97), Vol. 1, pp. PL1-PL6, 1997.
- [27] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k-means clustering algorithm," J. of the Royal Statistical Society, Series C (Applied Statistics), Vol.28, No.1, pp. 100-108, 1979.
- [28] B. Fritzke, "A growing neural gas network learns topologies," Advances in Neural Information Processing Systems, pp. 625-632, MIT Press, 1995.



Name:
Ryosuke Tanaka

Affiliation:
Human Mechatronics Systems, Graduate School of System Design, Tokyo Metropolitan University

Address:

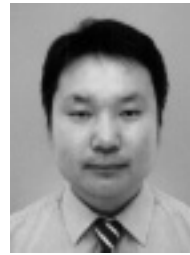
6-6 Asahigaoka, Hino, Tokyo 191-0065, Japan

Brief Biographical History:

2018 Received Bachelor of Engineering degree from Faculty of System Design, Tokyo Metropolitan University
2018- Master Course Student, Graduate School of System Design, Tokyo Metropolitan University

Main Works:

- R. Tanaka, J. Woo, and N. Kubota, "Cognitive Environment System by Joint Attention Behaviors and Relevance Theory for Robot Partners," 2018 IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC 2018), 2018.



Name:
Jinseok Woo

Affiliation:

Assistant Professor, Department of Mechanical Engineering, Faculty of Engineering, Tokyo University of Technology

Address:

1404-1 Katakura, Hachioji, Tokyo 192-0982, Japan

Brief Biographical History:

2011 Received Master of Engineering degree from Graduate School of System Design, Tokyo Metropolitan University
2017 Received Ph.D. degree from Graduate School of System Design, Tokyo Metropolitan University
2017-2019 Research Assistant Professor at Intelligent Mechanical Systems, Faculty of System Design, Tokyo Metropolitan University
2019- Assistant Professor at Department of Mechanical Engineering, Faculty of Engineering, Tokyo University of Technology

Main Works:

- J. Woo, J. Botzheim, and N. Kubota, "A Socially Interactive in Japanese Robot Partner Using Content-based Conversation System for Information Support," J. Adv. Comput. Intell. Inform., Vol.22, No.6, pp. 989-997, 2018.
- J. Woo, J. Botzheim, and N. Kubota, "System Integration for Cognitive Model of a Robot Partner," Intelligent Automation & Soft Computing (AUTOSOFT), pp. 1-14, 2017.

Membership in Academic Societies:

- The Institute of Electrical and Electronic Engineers (IEEE)
- The Japanese Association of Research on Care and Welfare (JARCW)



Name:
Naoyuki Kubota

Affiliation:

Professor, Graduate School of System Design, Tokyo Metropolitan University

Address:

6-6 Asahigaoka, Hino, Tokyo 191-0065, Japan

Brief Biographical History:

2005- Associate Professor, Department of System Design, Tokyo Metropolitan University
2007 Visiting Professor, University of Portsmouth
2009-2012 Visiting Professor, Seoul National University
2012- Professor, Graduate School of System Design, Tokyo Metropolitan University

Main Works:

- D. Tang, B. Yusuf, J. Botzheim, N. Kubota, and C. S. Chan, "A novel multimodal communication framework using robot partner for aging population," Expert Systems with Applications, Vol.42, Issue 9, pp. 4540-4555, 2015
- N. Kubota and Y. Toda, "Multi-modal Communication for Human-friendly Robot Partners in Informationally Structured Space," IEEE Trans. on Systems, Man, and Cybernetics – Part C, Vol.42, No.6, pp. 1142-1151, 2012.

Membership in Academic Societies:

- The Institute of Electrical and Electronic Engineers (IEEE)
- Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT)
- The Society of Instrument and Control Engineers (SICE)