https://doi.org/10.20965/jaciii.2007.p1274

Paper:

Joint Attention Between a Human Being and a Partner Robot Based on Computational Intelligence

Naoyuki Kubota*,, Toshiyuki Shimizu*, and Minoru Abe*****

*Dept. of System Design, Tokyo Metropolitan University
6-6 Asahigaoka, Hino, Tokyo 191-0065, Japan
E-mail: kubota@comp.metro-u.ac.jp, simizu-tosiyuki@ed.tmu.ac.jp
**SORST, Japan Science and Technology Agency (JST)
***EQUOS RESEARCH Co., Ltd. EV Laboratory
E-mail: i32561_abe@aisin-aw.co.jp

[Received December 14, 2005; accepted April 16, 2006]

Discussing joint attention between a human being and a partner robot to ensure human-friendly communication, we extract the direction of human attention based on facial direction in images, and propose control of the partner robot based on the direction of the extracted human face. We propose extraction for the direction in which a hand points, and discuss experimental results for partner robots using our proposal.

Keywords: intelligent robot, visual perception, human-friendly communication, genetic algorithms, neural network

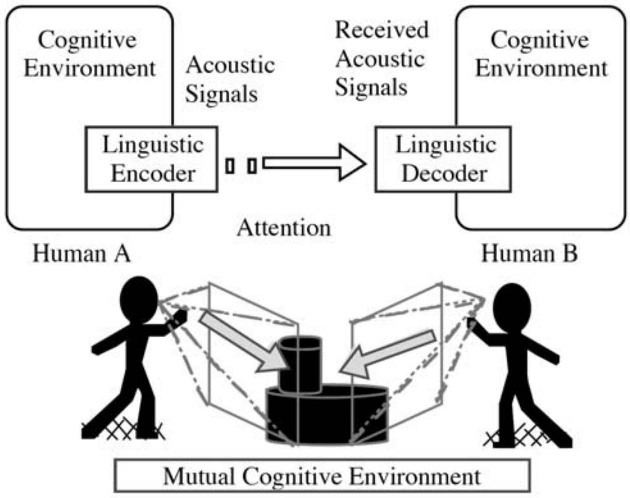


Fig. 1. Mutual cognitive environment in communication between humans.

1. Introduction

To realize human-friendly communication, a partner robot requires capabilities such as pattern clustering, pattern classification, voice recognition, situated utterance, and human-like actions. Social communication has been discussed in sociology, developmental psychology, semiotics, relevance theory, and embodied cognitive science [9].

In classical communication channel theory concerning a code model, a piece of information is transmitted as a symbol or signal based on an encoder and decoder. The symbol's meaning is shared between two systems if the two systems use the same encoder, decoder, protocol, and knowledge database. This explains communication from the viewpoint of information transmission, but people do not exactly share the same mechanism for interpretation and the same knowledge, so we discuss human communication from a different viewpoint.

In relevance theory helpful in discussing human communication [9], human thought is shared, rather than transmitted, between two people, each having own cognitive environment (**Fig. 1**). A person understands the meaning of an unknown word spoken by the other through communication because the person makes the symbol correspond to a percept, although they speak different languages. An important role of utterance and gesture is thus getting attention, and the utterance and gesture enlarge the cognitive environment of others. The shared cognitive environment is called a mutual cognitive environment. Joint attention plays the role of sharing cognitive environments.

Visual perception is the ability to detect light and to interpret it as perception [18]. Images of the world are not given to human beings, but are constructed by them. World construction is based on the search for patterns in objects and events. Pattern extraction is generally a conscious or unconscious search but the extraction of patterns by more than one person may be interactive and collaborative search. The attention of one person is decided based on that of others. Children use non-verbal communication such as gazing and proto-declarative pointing to share their attention toward an object or a third person with others, called joint attention skill [2, 3, 17]. Two people perceive objects requiring attention because the aim of joint attention is to communicate with a person through specific objects in the shared environment. Proto-declarative pointing is important in achieving joint attention, and hu-

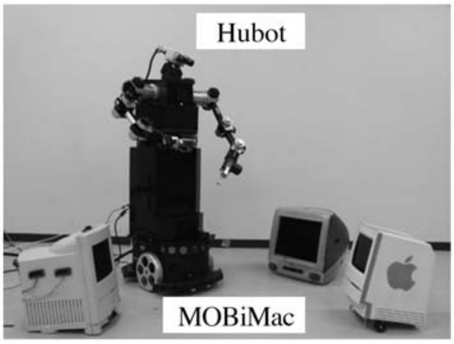


Fig. 2. Partner robots: Hubot and MOBiMacs.

man beings find it natural and effective to use gestures such as pointing to give a robot a specific task. A gesture is regarded as a symbolic action for communicating an intention to others, so the robot must recognize gestures.

Robots should also have a cognitive environment, and joint attention based on such a cognitive environment is very important for realizing social communication between human beings and robots. In discussing how to achieve the joint attention between a human and a robot based on human face and positioning recognition of the human arm, we propose extracting the direction of the human face by a steady-state genetic algorithm (SSGA) and a spiking neural network (SNN). We apply fuzzy motion control of a pan-tilt CCD camera. We propose a positioning recognition of a human arm by SSGA, and demonstrate in experiments in human face and gesture recognition for a partner robot the effectiveness of our proposal.

2. Partner Robots

We developed two types of partner robots – a human-like robot called Hubot [14] and a mobile PC called MOBiMac [15] – to realize social communication with a human (**Fig. 2**).

Hubot consists of a mobile base, a body, two arms with grippers, and a pan-tilt head. The robot has two CCD cameras, four line infrared sensors, a microphone, ultrasonic sensors, and touch sensors. Each CCD camera captures an image within a range of $-30°$ and $+30°$ in front of the robot. Two CPUs are used for sensing, motion control, and communication. The robot assumes human-like behavior. MOBiMac consists of two CPUs used for PC and robotic behavior. The robot has two servomotors, four ultrasonic sensors, four light sensors, a microphone, and a CCD camera.

Basic robot behavior involves visual tracking, map building, imitative learning, human classification, and voice recognition. Communication with a human being is conducted through utterances resulting from voice and human motion recognition [15]. Behavioral learning includes reinforcement learning through interaction with the environment, and imitative learning through interaction with human beings [14]. Behavioral interaction includes soccer and games with human beings. We focus on the visual perception of human face and positioning recognition of a human arm.

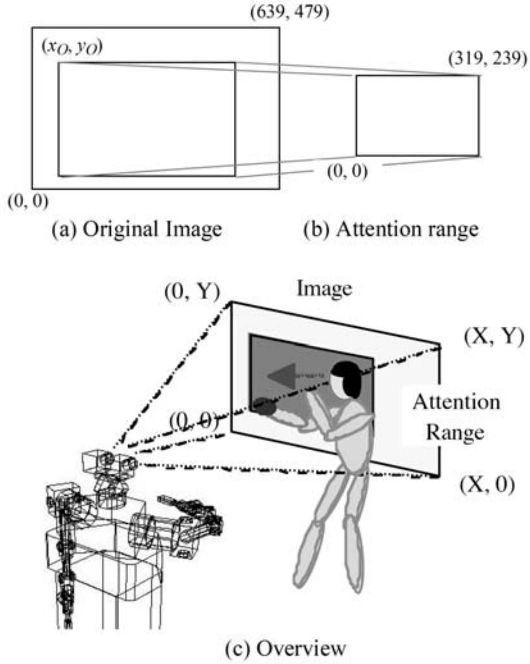


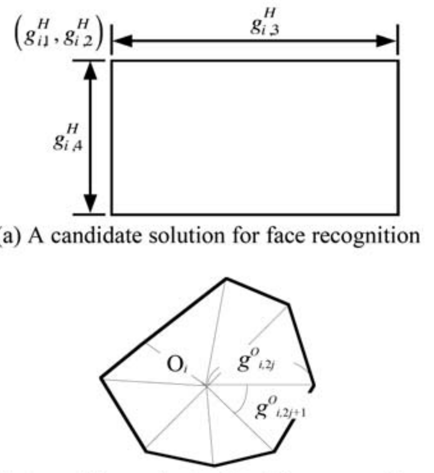
Fig. 3. Human face detection for joint attention.

3. Joint Attention

3.1. Human Face Detection

Human face recognition consists of differential extraction, human face detection, and extraction of facial direction. An image of RGB color space is taken by a CCD camera with the partner robot. Because image processing takes time and computational cost, full-image size of image processing of all images is not practical. The original image is 640×480, reduced to 320×240 as an attention range based on reduction level ($1 \leq RL \leq 2$) and the origin (x_O, y_O) of the attention range (**Fig. 3**). If the reduction level is 1, the same image resolution is extracted from the original image. Otherwise, each pixel in the attention range is interpolated based on the four surrounding pixels based on the reduction level.

Colors corresponding to hair and skin are extracted using thresholds. Candidate regions including the human face and hair color are detected using a steady-state genetic algorithm (SSGA) based on template matching [4, 15]. A candidate solution of the template used for detecting a human as the target object consists of numerical parameters of $g_{i,1}^H$, $g_{i,1}^H$, $g_{i,3}^H$, and $g_{i,4}^H$ where $(g_{i,1}^H, g_{i,2}^H)$ is the point of the upper left corner of the template; $g_{i,3}^H$ is the width; $g_{i,4}^H$ is the height (**Fig. 4(a)**). The number of individuals is G. To speed up the search, we use the



(a) A candidate solution for face recognition

(b) A candidate solution for object recognition

Fig. 4. Genotype representation in SSGA.

further reduced image size for detecting a moving object, that is, recognizable as a human being. The image size used in differential extraction is 20×15. The center of gravity (COG) in differential extraction is used for updating candidate solutions. In SSGA, few existing solutions are replaced with the candidate generated by genetic operators in each generation. The worst candidate solution is eliminated ("Delete least fitness" selection), replaced by the candidate solution generated by crossover and mutation. The fitness value is calculated as follows:

$$f_i^H = C^H_{\text{Skin}} + C^H_{\text{Hair}} + \eta_1^H \cdot C^H_{\text{Skin}} \cdot C^H_{\text{Hair}} - \eta_2^H \cdot C^H_{\text{Other}} \quad . \quad (1)$$

where C^H_{Skin}, C^H_{Hair}, and C^H_{Other} indicate the numbers of pixels of colors corresponding to human skin, human hair, and other colors. η_1^H and η_2^H are coefficients. Therefore, this problem results in a maximization problem. We use elitist crossover and adaptive mutation. Elitist crossover randomly selects one individual and generates an individual by combining genetic information from the randomly selected individual and the best individual. Next, adaptive mutation is conducted for the generated individual:

$$g^H_{i,j} \leftarrow g_{i,j} + \left(\alpha_j^H \cdot \frac{f^H_{\max} - f^H_i}{f^H_{\max} - f^H_{\min}} + \beta_j^H\right) \cdot N(0,1) \quad (2)$$

where f_i^H is the fitness value of the i-th individual; $f^H_{\max}$ and $f^H_{\min}$ are the maximum and minimum of fitness values in the population; $N(0,1)$ indicates a normal random value; α_j^H and β_j^H are the coefficient and offset. In adaptive mutation, the variance of the normal random number is changed based on fitness values of the population. By using SSGA, the robot roughly detects a human face candidate.

Facial direction is roughly extracted using the relative positioning of the hair and face, i.e., the relative position of COG of areas corresponding to the hair and face. The relative positions of COG versus the central position of the detected face region are used as inputs to spiking neurons for extracting the facial direction. We apply a SNN to extract the direction of the detected human face. SNNs have been used in memorizing spatial and temporal contexts [5, 6]. A SNN is considered as one of the artificial NNs imitating the dynamics introduced in the ignition phenomenon of a cell and the propagation mechanism of the pulse between cells. We use a simple spike response model to reduce computational cost. Internal state $h_i(t)$ is calculated as follows:

$$h_i(t) = \tanh\left(h_i^{\text{syn}}(t) + h_i^{\text{ext}}(t) + h_i^{\text{ref}}(t)\right). \quad . \quad . \quad . \quad (3)$$

$h_i^{\text{ext}}(t)$ is the input to the i-th neuron from the external environment and $h_i^{\text{syn}}(t)$ including output pulses from other neurons is calculated as follows:

$$h_i^{\text{syn}}(t) = \gamma^{\text{syn}} \cdot h_i(t-1) + \sum_{j=1, j\neq i}^{N} w_{j,i} \cdot p_j(t-1). \quad (4)$$

$h_i^{\text{ref}}(t)$ indicates the refractoriness of the neuron; $w_{j,i}$ is a weight coefficient from the j-th to i-th neuron; $p_j(t)$ is the output of the j-th neuron at the discrete time t; N is the number of neurons; and γ^{syn} is a discount rate. When the neuron is fired, R is subtracted from the refractoriness value as follows:

$$h_i^{\text{ref}}(t) = \begin{cases} \gamma^{\text{ref}} \cdot h_i^{\text{ref}}(t-1) - R & \text{if } p_i(t-1) = 1 \\ \gamma^{\text{ref}} \cdot h_i^{\text{ref}}(t-1) & \text{otherwise} \end{cases} \quad (5)$$

where γ^{ref} is a discount rate. When the internal state of the i-th neuron is larger than the predefined threshold, a pulse is output as follows:

$$p_i(t) = \begin{cases} 1 & \text{if } h_i^{\text{ref}}(t) \leq q_i \\ 0 & \text{otherwise} \end{cases} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where q_i is a threshold for firing.

We use two layers of SNN, the first layer consisting of sensor neurons for extracting the facial direction and the second consisting of motor neurons for control of the pan-tilt CCD camera, so the robot's head moves in the direction to which the person pays attention.

3.2. Object Recognition

We focus on color-based object and shape recognition with SSGA based on template matching. The shape of a candidate template is generated by the SSGA. We proposed extracting a hand by using SSGA, and used an octagonal template of the angle fixed at 45° [14]. To improve shape recognition, we propose an octagonal template with variable angles. **Fig. 4(b)** shows a candidate template used for detecting a target where the j-th point $g^O_{i,j}$ of the i-th template is represented by $(g^O_{i,1} + g^O_{i,j}\cos(g^O_{i,j+1}), g^O_{i,2} + g^O_{i,j}\sin(g^O_{i,j+1}))$, $i = 1, 2, \ldots, n$, $j = 1, \ldots, 2 \times m + 2$; O_i $(= (g^O_{i,1}, g^O_{i,2}))$ is the center of a candidate template on the image; n and m are the number of candidate templates and search points used in a template. A candidate template consists of numerical parameters of $(g^O_{i,1}, g^O_{i,2}, \cdots, g^O_{i,2m+2})$. We propose a local search for fitting a template to the area of the target object. The central position is slightly updated by calculating the COG of search points of the octagon. Search points and angles are updated by hill-climbing. If the color of a search point randomly extended toward the outside corresponds to the target color, the generated search point is replaced with the previous search point. This local search is conducted only for the template candidate with a high fitness value, calculated as follows:

$$f_i^O = C_{\text{Target}}^O - \eta_1^O \cdot C_{\text{Other}}^O \quad \dots\dots\dots\dots \quad (7)$$

where η_1^O is a coefficient for penalty; C_{Target}^O and C_{Other}^O indicate the numbers of pixels of colors corresponding to a target and other colors included in the template. The target color is selected based on the pixel color occupying most of the template candidate, so the largest area of a single color is extracted in the reduced color space of the image. One iteration consists of selection, crossover, and mutation. The iteration of SSGA is repeated until the termination condition is satisfied. In general, the maximum number of evaluations or generations is used as the termination condition.

We apply a K-means algorithm for clustering candidate templates to find several objects simultaneously. The K-means algorithm is widely used in iterative descent clustering [8]. Inputs to the K-means algorithm are the central position of templates candidates; $\mathbf{v}_j$ $(= (g_{j,1}^O, g_{j,2}^O))$, $j = 1, 2, \dots, n)$. When the reference vector of the i-th cluster is represented by $\mathbf{r}_i$, the Euclidian distance between the j-th input vector and the i-th reference vector is defined as

$$d_{j,i} = ||\mathbf{v}_j - \mathbf{r}_i|| \quad \dots\dots\dots\dots \quad (8)$$

where $\mathbf{r}_i = (r_{i,1}, r_{i,2})$ and the number of reference vectors (output units) is l. The reference vector minimizing distance $d_{j,i}$ is selected by

$$c_j = \arg\min_i \{||\mathbf{v}_j - \mathbf{r}_i||\} \quad \dots\dots\dots\dots \quad (9)$$

where c_j is the cluster number that the j-th input belongs to. After selecting the reference vector nearest to each input, the i-th reference vector is updated by the average of inputs belonging to the i-th cluster. If updating is not conducted in clustering, updating is finished. Crossover and selection are conducted within template candidates of each cluster, so SSGA attempts to find different objects within each cluster based on the spatial distribution of objects in the image.

Spiking neurons are applied for shape recognition of detected color objects. We use three sensor neurons for extracting circles, triangles, and rectangles. The number of acute angles in a template candidate is used for sensor neurons. If the sensor neuron corresponding to each shape is fired, the spike output is transmitted to the utterance system and map building for a mutual cognitive environment [19].

The facial direction is used for restricting the search area for image processing and for motion control of the CCD camera. The attention area is updated based on the facial direction. If the robot detects a person in the attention area, it attempts to physically interact with the person. If the robot detects objects, it attempts to visually track the target object. In this way, the robot archives joint attention with the human.

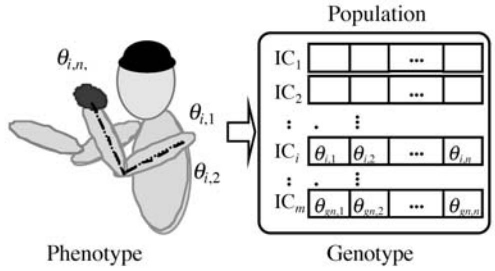


Fig. 5. Genotype representation for positioning extraction of the arm.

3.3. Human Arm Recognition

The 3D positioning of the arm is mapped into a 2D image when a camera is used, so we must solve the problem of inverting from the 2D image to 3D positioning. Configuration $\boldsymbol{\theta}$ of the arm is expressed by a set of joint angles:

$$\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_n)^T \in R^n \quad \dots\dots\dots \quad (10)$$

where n denotes the DOF of the arm. In this paper, the DOF of the arm is assumed to be 5 ($n = 5$). The position of the hand is expressed as $\boldsymbol{P} = (p_x \; p_y \; p_z)^T$ on the base frame. SSGA is applied to detect joint angles corresponding to the arm positioning in the image.

An individual is consists of all joint variables (**Fig. 5**). Initialization updates the population based on the previous best configuration to the next image. The j-th joint angle of i-th configuration $\theta_{i,j}$, represented as a real number, is generated as follows ($i = 1, 2, ..., g_n$):

$$\theta_{i,j} \leftarrow \theta_j^* + \beta_j^I \cdot N(0,1) \quad \dots\dots\dots\dots \quad (11)$$

where θ_j^* is the previous best joint angle; β_j^I is a coefficient for the j-th joint angle. We use a following fitness function:

$$f_i^A = f_{\text{Distance}}^A + \eta_1^A \cdot f_{\text{Angle}}^A + \eta_2^A \cdot f_{\text{Color}}^A \quad \dots\dots \quad (12)$$

where η_1^A and η_2^A are weight coefficients; f_{Distance}^A, denotes the distance between the hand position and the target point; f_{Angle}^A denotes the sum of squares of the difference between each joint angle between two configurations of t and $t-1$; f_{Color}^A denotes the number of pixels corresponding to skin color. A selection removes the worst individual from the current population, followed by elitist crossover. The worst individual is thus replaced by the individual generated by the elitist crossover. We use adaptive mutation, and the robot extracts the human arm positioning.

4. Experiments

In experiments, the image size for human detection is 320×240 and that for object recognition is 160×120 because object recognition does not require high image resolution. The population size of SSGA for human face detection is 30, for object recognition is 100, and for arm

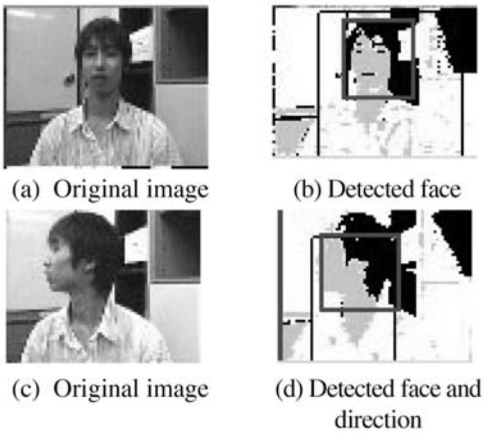


Fig. 6. Recognition result for facial direction.

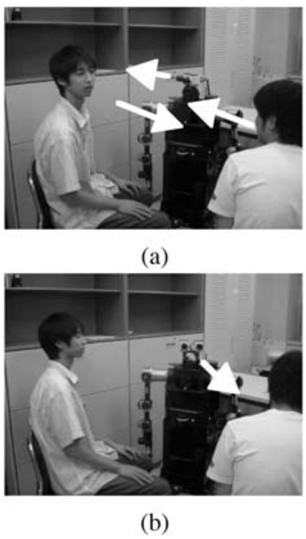


Fig. 7. Head movement based on facial direction.

positioning recognition is 200.

Figure 6 shows experimental results of the extraction of facial direction for constructing joint attention by Hubot. The image for human detection by Hubot is 160×120 due to the computational performance of the vision system consisting of two CCD cameras. **Fig. 7** shows the positioning of two people and the robot. In this experiment, the robot attempts to extract the direction of the eyes in a conversation between two people. **Figs. 6(a)** and **(c)** show original images, and **Figs. 6(b)** and **(d)** extraction results in joint attention between them. Red rectangles in **Figs. 6(b)** and **(d)** indicate the extracted human face. The red bar of the edge in **Fig. 6(d)** indicates the extracted direction of the human face. The robot concludes that the person on the right is looking at the person on the left, and robot moves the pan-tilt head to the right. After searching for persons or objects, the robot stops in front of the face of the other person, and extracts the facial direction of the person at left, determining based on joint attention that a conversation is going on between the two people (**Fig. 7**).

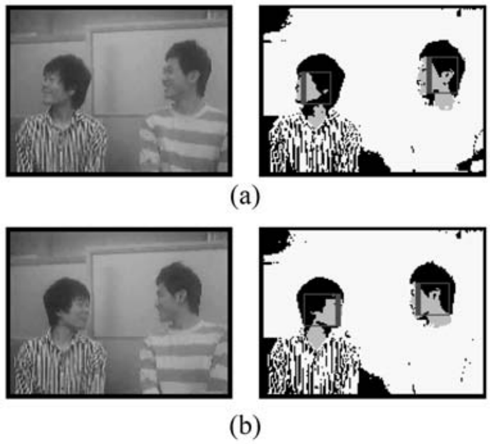


Fig. 8. Extraction of facial direction of two people.

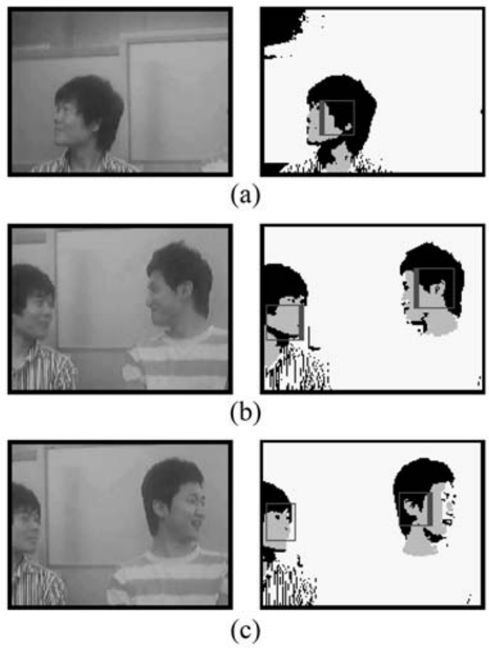


Fig. 9. Extraction of facial direction of two people based on the attention range control.

Figure 8 shows results of extraction of the facial direction of two people in the same image by MOBiMac where the reduction level is 2.0. A red thick bar is shown when the facial direction is recognized by spiking neurons, showing that our proposed method extracts the direction of faces successfully.

Figure 9 shows results of extraction of facial direction of two people in the same image based on the attention range control by MOBiMac. Since the reduction level for constructing the attention range is 1.3, it is difficult for the robot to capture two people in the attention range, so it attempts to construct the attention range based on the extracted facial direction of the two people. The origin of the attention range is updated by SNN output. In **Fig. 9(a)**, the attention range is constructed on the left side of the original image, and in **Fig. 9(b)**, it is constructed on

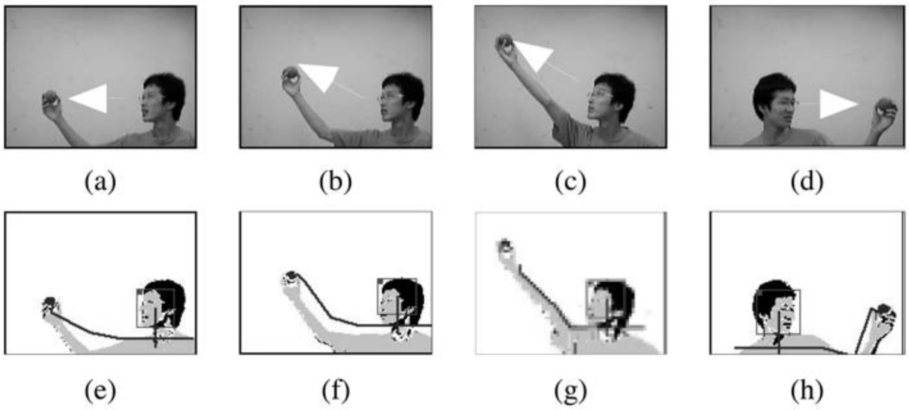


Fig. 11. Image processing for extracting the human arm posture.

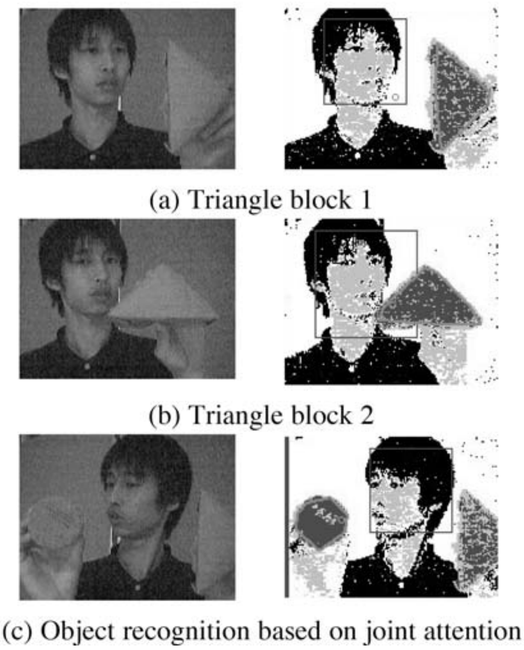


Fig. 10. Recognition result of facial direction.

the right. In **Fig. 9(c)**, the attention range is constructed midway between two people.

Figure 10 shows results of object recognition for a triangle block from different angles. Although we used an octagonal template for object recognition, the triangle is extracted. **Fig. 10(c)** shows object recognition based on joint attention between the human and robot. After the robot extracts the human facial direction, the search of objective recognition is restricted to the left.

Figure 11 shows positioning recognition. The extracted facial direction is used to search for object and human hand detection. When the person sees right hand in **Figs. 11(a)**, **(b)**, and **(c)**, he sees the left hand in **Fig. 11(d)**. Based on the human face, the arm suitable for positioning recognition is selected in the search, so the robot uses the direction of the face in the search based on the joint attention.

5. Summary

We have proposed achieving joint attention required in social communication, applying steady-state genetic algorithms and spiking neural networks for human detection, object recognition, and arm positioning extraction. Preliminary results in visual perception show that our proposed method extracts the direction of human attention to archive joint attention with a human, but we did not discuss how to realize the social communication based on our proposal. We therefore must combine voice recognition and visual perception to share the cognitive environment. In other projected work, we will discuss how to solve the symbol grounding problem by having the robot learns through interaction with a person.

References:

[1] R. A. Brooks, "Cambrian Intelligence," The MIT Press, 1999.
[2] R. Pfeifer and C. Scheier, "Understanding Intelligence," The MIT Press, 1999.
[3] B. Scassellati, "Investigating Models of Social Development using a Humanoid Robot," Robots and Biology: Developing Connections, 1998.
[4] G. Syswerda, "A Study of Reproduction in Generational and Steady-State Genetic Algorithms," Foundations of Genetic Algorithms, San Mateo: Morgan Kaufmann Publishers, Inc., 1991.
[5] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, "Neuro-Fuzzy and Soft Computing," Prentice-Hall, Inc., 1997.
[6] W. Maass and C. M. Bishop, "Pulsed Neural Networks," The MIT Press, 1999.
[7] T. Kohonen, "Self-Organization and Associative Memory," Springer, 1984.
[8] T. Hastie, R. Tibshirani, and J. Friedman, "The Elements of Statistical Learning," Springer-Verlag, 2001.
[9] D. Sperber and D. Wilson, "Relevance - Communication and Cognition," Blackwell Publishing Ltd., 1995.
[10] K. Morikawa, S. Agarwal, C. Elkan, and G.Cottrell, "A Taxonomy of Computational and Social Learning," Proc. of Workshop on Developmental Embodied Cognition, 2001.
[11] G. Rizzolatti, and M. A. Arbib, "Language within our grasp," Trends in Neuroscience, Vol.21, pp. 188-194, 1998.
[12] T. Fukuda and N. Kubota, "An Intelligent Robotic System Based on A Fuzzy Approach," Proc. of IEEE, Vol.87, No.9, pp. 1448-1470, 1999.
[13] N. Kubota, D. Hisajima, F. Kojima, and T.Fukuda, "Fuzzy and Neural Computing for Communication of A Partner Robot," Journal of Multiple-Valued Logic and Soft-Computing, Vol.9, No.2 pp. 221-239, 2003.
[14] N. Kubota and K. Tomoda, "Behavior Coordination of A Partner Robot based on Imitation," Proc. of 2nd Int. Conf. on Autonomous Robots and Agents, pp. 164-169, 2004.

[15] N. Kubota and K. Nishida, "Fuzzy Computing for Communication of A Partner Robot Based on Imitation," Proc. (CD-ROM) of 2005 IEEE Int. Conf. on Robotics and Automation, pp. 4391-4396, 2005.
[16] M. Merleau-Ponty, "Phenomenology of Perception," Trans. Colin Smith. New York: Humanities Press, 1962.
[17] B. Robins, P. Dickerson, P. Stribling, and K. Dautenhahn, "Robot-Mediated Joint Attention in Children with Autism," Interaction Studies Vol.5, No.2, pp. 161-198, 2004.
[18] H. H. Bulthoff, S. W. Lee, T. A. Poggio, and C. Wallraven, "Biologically Motivated Computer Vision," Springer-Verlag, 2002.
[19] N. Kubota and K. Nishida, "Development of Internal Models for Communication of A Partner Robot Based on Computational Intelligence," Proc. of 6th Int. Symposium on Advanced Intelligent Systems, pp. 577-582, 2005.

Name:
Naoyuki Kubota

Affiliation:
1. Associate Professor, Dept. of System Design, Tokyo Metropolitan University, Japan
2. PRESTO, Japan Science and Technology Agency (JST)

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

Brief Biographical History:
1997 Joined with the Dept. of Mechanical Engineering at Osaka Institute of Technology
1999 Joined as an Associate Professor at Dept. of Human and Artificial Intelligent Systems at University of Fukui
April, 2004 Joined as an Associate Professor at Dept. of System Design, Tokyo Metropolitan University

Main Works:
• N. Kubota and K. Nishida, "Cooperative Perceptual Systems for Partner Robots Based on Sensor Network," Int. Journal of Computer Science and Network Security (IJCSNS), Vol.6, No.11, pp. 19-28, 2006.
• N. Kubota, Y. Nojima, F. Kojima, and T. Fukuda, "Multiple fuzzy state-value functions for human evaluation through interactive trajectory planning of a partner robot," Soft Computing, Vol.10, No.10, pp. 891-901, 2006.
• N. Kubota, "Multi-Objective Design of Neuro-Fuzzy Controllers for Robot Behavior Coordination," Y. Jin (Ed.), Multi-Objective Machine Learning, Springer, pp. 557-584, 2006.

Membership in Academic Societies:
• Robotics Society of Japan (RSJ)
• Japan Society of Mechanical Engineers (JSME)
• Society of Instrument and Control Engineers (SICE)
• Institute of Electrical and Electronic Engineers (IEEE)

Name:
Toshiyuki Shimizu

Affiliation:
Dept. of System Design, Tokyo Metropolitan University, Japan

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

Brief Biographical History:
2006 Received an M.E. degree from the Department of Precision Engineering, Tokyo Metropolitan University

Main Works:
• N. Kubota and T. Shimizu, "Modular Fuzzy Neural Networks for Imitative Learning of A Partner Robot," Proc. of 2006 IEEE World Congress on Computational Intelligence (WCCI 06), pp. 8179-8186, 2006.

Membership in Academic Societies:
• Student Member, Institute of Electrical and Electronic Engineers (IEEE)
• Student Member, Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT)

Name:
Minoru Abe

Affiliation:
EQUOS RESEARCH Co., Ltd. EV Laboratory

Address:
Nakato Marunouchi Building, 3-17-6 Marunouchi, Naka-ku, Nagoya, Aichi 460-0002, Japan

Brief Biographical History:
Apr., 2004-Mar., 2006 Department of Human and Artificial Intelligent System University of Fukui
Apr., 2006-Jun., 2006 AISIN AW Co., Ltd.
Jul., 2006- EQUOS RESEARCH Co., Ltd.

Main Works:
• "Interactive Learning for A Partner Robot Based on Cyclic Gestures," Proc. (CD-ROM) of The 5th Int. Conf. on Simulated Evolution And Learning, SWA-1, 2004.
• "Computational Intelligence for Cyclic Gestures Recognition of A Partner Robot," Proc. of Ninth Int. Conf. on Knowledge-Based Intelligent Information & Engineering Systems, pp. 650-656, 2005.
• "Gestures Recognition for A Partner Robot Based on Spiking Neural Network," Proc. of Daeduck Int. Conf. on Human-Centered Advance Intelligent Technology, p. 1, 2005.

Membership in Academic Societies:
• Society of Automotive Engineers of Japan, Inc. (JSAE)