Remote Monitoring and Universal Remote Control Based on iPhone in Informationally Structured Space

Naoyuki Kubota, Rikako Komatsu,
Tokyo Metropolitan University
6-6 Asahigaoka, Hino, Tokyo 191-0065, Japan

Beom Hee Lee
Seoul National University
San 56-1, Shillim-Dong, GwanAk-Ku, Seoul 151-742, Kore

Abstract

This paper proposes a monitoring and remote control system based on human localization in informationally structured space using a sensor network. First, we explain informationally structured space, robot partners, remote control system for home appliances, and sensor networks for human motion measurement developed in this study. Next, we apply a spiking neural network to extract a person from the measured data by the sensor network. Furthermore, we propose a learning method of spiking neural network based on the time series of measured data. Finally, we discuss the effectiveness of the proposed methods through experimental results in a living room.

Keywords: Monitoring, Remote Control, Human Localization, Sensor Networks, Robot Partners, Neural Networks

1. INTRODUCTION

Recently, many home appliances such as audio players television, air-conditioner, lights, and fans have been controlled by infrared remote controllers. As a result, there are many remote controllers in a house. It is troublesome for elderly people to use many remote controllers in a living room, because each remote control has the different layout and functions of buttons. Furthermore, elderly people sometimes forget to turn home appliances off. Therefore, the monitoring system of states inside the house is helpful for them. On the other hand, various types of personal data assistant (PDA) devices, personal organizers, and smart phones have been developed to access the personal information and Internet information until now. Such a device can be used to control home appliances. However, the device is unfamiliar to elderly people, and it takes much time to choose the menu of the target home appliance from the candidates, as the number of home appliances increases in a house. Furthermore, the remote monitoring of elderly people living alone in a house is essential for their family. We proposed the concept of informationally structured space [10] (Fig.1). The environment surrounding people and robots should have a structured platform for gathering, storing, transforming, and providing information. The structuralization of informationally structured space realizes the quick update and access of valuable and useful information for users. Furthermore, we should consider the accessibility to required information, especially, human interface is very important to use devices [1,2].

In this paper, we propose a universal remote control system of home appliances in the informationally structured space. We use Apple iPhone for the device for remote control of home appliance, because iPhone can provide the multi-modal communication interface with users [20]. The iPhone can estimate the posture and direction of the device itself by the internal compass and accelerometer. In order to use iPhone as a remote controller, the proposed system must estimate the location of the user and the aim of the user. Therefore, we discuss the on-line estimation method of human location in the informationally structured space based on sensor networks. Next, we apply a spiking neural network [11-13] to localize human position, and to learn temporal relationship of behaviors based on the firing patterns. Finally, we show experimental results, and discuss the effectiveness of the proposed method.

This paper is organized as follows. Section 2 explains the robot partners, data flow in the informationally structured space, remote control for home appliances, spiking neural network for human localization. Section 3 shows experimental results of the proposed method, and Section 4 summarizes the paper, and discusses the future vision of robot partners.

2. INFORMATIONALLY STRUCTURED SPACE

Robot Partners

We can use three different types of robot partners from the interactive point of view (Fig.2). One is a physical robot partner. We can interact with the physical robot partner by using multi-modal communication like a human. The next one is a pocket robot partner. The pocket robot partner has no mobile mechanism, but we can easily bring it everywhere and can interact with the robot partner by touch and physical interface. The last one is a virtual robot partner. The virtual robot partner is in the virtual space in the computer, but we can immerse into the virtual space, and interact with it through the virtual person or robot. The interaction style of these three types of robot partners is different, but they share the same personal database and interaction logs, and can interact with the person based on the same interaction rules independent from the style of interfaces.

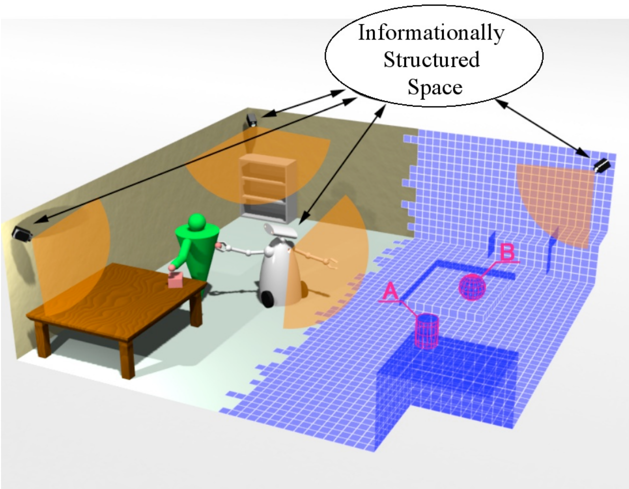


Fig.1. The concept of informationally structured space

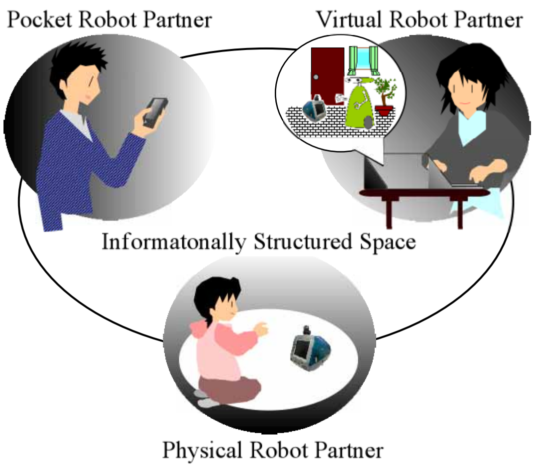


Fig.2. Interaction with robot partners

We developed two types of physical robot partners; a mobile PC called MOBiMac and a human-like robot called Hubot in order to realize the social communication with a human [14-17]. Each robot has two CPUs and many sensors such as CCD camera, microphone, and ultrasonic sensors. Therefore, the robots can conduct image processing, voice recognition, target tracing, collision avoidance, map building, imitative learning, and others [5,14-20]. The basic behavior modes of these robots are human tracking, human communication, behavioral learning, and behavioral interaction. The communication with a person is performed by utterance as the result of voice recognition and gestures as the result of human motion recognition. The robot integrates the above behaviors according to the environmental conditions based on the multi-objective behavior coordination. The multi-objective behavior coordination integrates outputs of several behaviors according to the time-series of perceptual information.

We have used iPhone and iPod touch as a pocket robot partner, because we can easily use the touch interface, accelerometer, compass, and GPS in the program development. This device can be used for tele-operation of robots and remote monitoring in addition to personal data assistance.

The basic capabilities common to physical, pocket, and virtual robot partners are human recognition, object recognition, mining of personal data, and learning of interaction patterns based on image processing and voice recognition. The data used for these capabilities are stored in the informationally structured space, and a robot partner can access and update the data through the wireless network.

Data Flow in Informationally Structured Space

Figure 3 shows the data flow in the developed system based on informationally structured space. The developed system is divided into four main components; (1) database management server PC, (2) physical robot partners, (3) environmental systems, and (4) pocket robot partners as human interface systems. The environmental system is based on a wireless sensor network composed of sensors equipped with wall, floor, ceiling, furniture, and home appliances. These sensors measure the environmental data and human motions. The measured data are transmitted to the database server PC, and then feature extraction is performed. Each robot partner can receive the environmental information from the database server PC, and serves as a partner to the person. Furthermore, the user interface system is used for the person to access the environmental information through pocket robot partners. Here the learning infrared controller is used for the control of home appliances such as air-conditioner, TV, and lights.

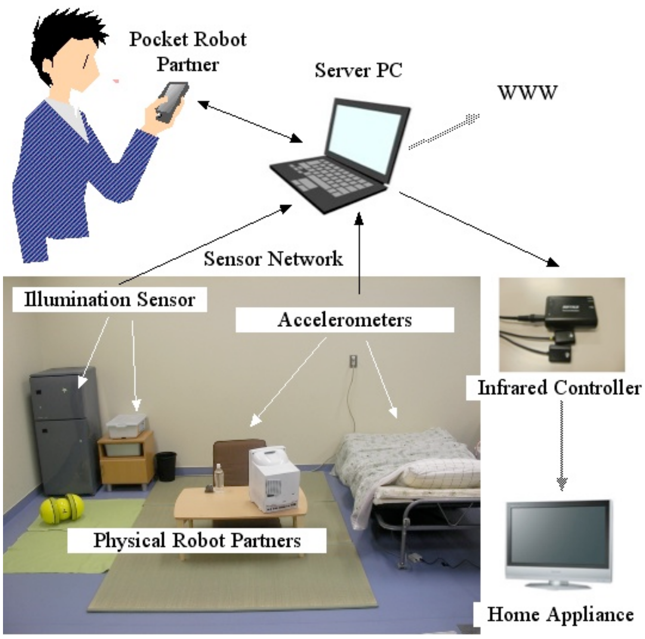
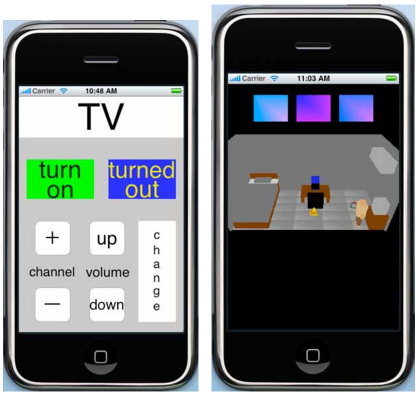


Fig.3. Data flow in informationally structured space



(a) CUI mode (b) GUI mode

Fig.4. Interface modes of Universal remote controller based on pocket robot partner

Fig.5. Universal remote controller for home appliance

Remote Control of Home Appliances

In general, there are many remote controllers for home appliances in a house, and the layout of buttons in controllers is different. Therefore, one universal remote controller for controlling all home appliances is reasonable and useful for people. The pocket robot partner is used for the multi-modal communication interface to control home appliances. iPhone provide a person with human interfaces for voice, tough, and gesture. Figure 4 shows two modes of touch interface based on character-based user interface (CUI) and graphical user interface (GUI). Originally, the term of CUI is used as the interface done by only texts, but we use the term of CUI in this paper.

In the CUI mode, the menu of controller is composed of the name of home appliance, control of power supply and adjustment from the top to bottom. If a person turns the iPhone to a home appliance, the menu of home appliance is automatically changed. The color and size of fonts and buttons are designed according to the results of questionnaire (see Fig.4 (a)). Because the position of the person is localized by the sensor network, the direction of the iPhone can be detected by the compass and accelerometer in the iPhone (Fig.5). If a home appliance is turned on, the menu is changed from the combination of “turn on” and “turned out” to that of “turned on” and “turn out”. In general, the switch for power supply is only one in a standard remote controller, but elderly people sometimes push the button twice owing to the shaking of a finger. Therefore, we divide the switch for power supply into two buttons.

Furthermore, in the GUI mode, the user can directly touch a home appliance in the simulator to turn on/off its corresponding home appliance. If the home appliance is turned on, its color in the simulator is brightened. In this way, the user can easily understand the state of home appliance in the GUI mode. However, it is difficult to adjust the state of home appliances, e.g., the temperature of air, and the channel and volume of TV in the GUI mode. Therefore, we intend to incorporate the small menu for the adjustment of home appliances in the GUI mode like augmented reality environments as a future work.

Sensor Network for Human Motion Measurement

We use a 3D distance image sensor, SR-3000 (Fig.6 (a)) developed by MESA Corporation. SR-3000 is a range camera for measuring 3-dimensional distance up to 7.5 [m] based on time-of-flight principal by using infrared light source, and outputs the measured data by a USB2.0 interface. The output data is composed of distance data as a Cartesian axis and luminance data with a spatial QCIF resolution (176 * 144 pixels). Furthermore, we use 3-axis Jyro to measure human behaviors. The wireless sensor device called Sun SPOT (Sun Small Programmable Object Technology) with 3-axis Jyro is included in a block (Fig.6 (b)). Sun SPOT is small, wireless, battery-powered device powered by a specially designed small-footprint Java virtual machine, called Squawk, that can host multiple applications concurrently, and requires no underlying operating system.

(a) 3D distance image sensor (b) Wireless sensor device

Fig.6. Measurement system for human behaviors

Spiking Neural Network for Human Localization

Various types of artificial neural networks have been proposed to realize clustering, classification, nonlinear mapping, and control [11-13, 20-22]. Basically, artificial neural networks are classified into pulse-coded neural networks and rate-coded neural networks from the viewpoint of abstraction level. A pulse-coded neural network approximates the dynamics with the ignition phenomenon of a neuron, and the propagation mechanism of the pulse between neurons. Hodgkin-Huxley model is one of the classic neuronal spiking models with four differential equations. An integrate-and-fire model with a first-order linear differential equation is known as a neuron model of a higher abstraction level. A spike response model is slightly more general than the integrate-and-fire model, because the spike response model can choose kernels arbitrarily.

One important feature of pulse-coded neural networks is the capability of temporal coding. In fact, various types of spiking neural networks (SNNs) have been applied for memorizing spatial and temporal context. We use a simple spike response model to reduce the computational cost. The internal state $h_i(t)$ is calculated as follows;

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

Here hyperbolic tangent is used to avoid the bursting of neuronal fires, $h_i^{ext}(t)$ is the input to the ith neuron from the external environment, and $h_i^{syn}(t)$ including the output pulses from other neurons is calculated by,

$$h_i^{syn}(t) = \gamma^{syn} \cdot h_i(t-1) + \sum_{j=1, j\neq i}^{N} w_{j,i} \cdot h_j^{PSP}(t) \quad (2)$$

Furthermore, $h_i^{ref}(t)$ indicates the refractoriness factor of the neuron; $w_{j,i}$ is the parameter of a weight coefficient from the jth to ith neuron; $h_j^{PSP}(t)$ is the presynaptic action potential (PSP) approximately transmitted from the jth neuron at the discrete time t; N is the number of neurons; γ^{syn} is a temporal discount rate. When the internal action potential of the ith neuron is larger than the predefined threshold, a pulse is outputted as follows;

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

where q_i is a threshold for firing. Furthermore, R is subtracted from the refractoriness value in the following,

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

where γ^{ref} is a discount rate and R>0.

The presynaptic spike output is transmitted to the connected neuron according to PSP with the weight connection. The PSP is calculated as follows;

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

where γ^{PSP} is the discount rate (0< γ^{PSP} <1.0). Therefore, the postsynaptic action potential is excitatory if the weight parameter, $w_{j,i}$ is positive. If the condition $h_j^{PSP}(t-1) < h_i^{PSP}(t)$ is satisfied, the weight parameter is trained based on the temporal Hebbian learning rule as follows,

$$w_{j,i} \leftarrow \tanh\left(\gamma^{wgt} \cdot w_{j,i} + \xi^{wgt} \cdot h_j^{PSP}(t-1) \cdot h_i^{PSP}(t)\right) \quad (6)$$

where γ^{wht} is a discount rate and ξ^{wgt} is a learning rate.

We apply SNN to the human localization based on measured data of the sensor networks. Basically, each furniture or equipment is attached with a sensor. If the measured value is changed large, then the difference from the base value is used as the inputs to a spiking neuron in the following;

$$h_i^{ext}(t) = \min\left\{\left(v_i(t) - V_i\right)^2, 1\right\} \quad (7)$$

where $v_i(t)$ is the measured values at t and V_i is the base value of the ith sensor. Here, the input value to the spiking neuron is fuzzified. Furthermore, the base value is updated in case of accelerometer equipped with the movable objects such as chair or bad,

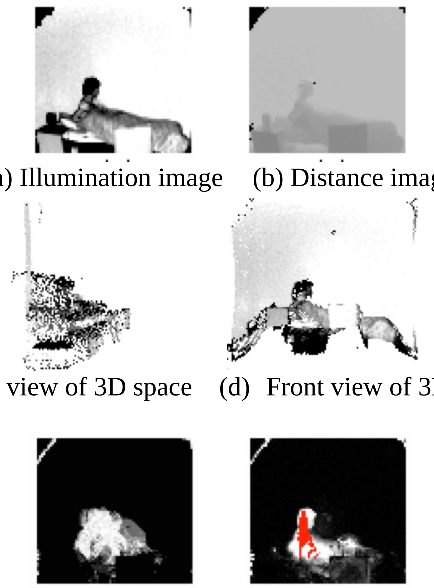
$$V_i \leftarrow (1-\xi^{sen}) \cdot V_i + \xi^{sen} \cdot v_i(t) \quad (8)$$

where ξ^{sen} is the learning rate to adjust the base value to the current situation of the movable objects. If the neuron is fired, this means that a person uses or moves its corresponding furniture. In this way, the firing pattern indicates the time-series of human position or behavior in the room. If the position of each sensor is not localized, the simultaneous firing of sensor neuron and human walking indicates the high possibility that the person uses the furniture or equipment. Based on this discussion, we can assume the position of the person as that of the ith furniture or equipment, ($X_{i,1}^S$, $X_{i,2}^S$).

$$X_{i,k}^S \leftarrow (1-\alpha^S) X_{i,k}^S + \alpha^S \cdot \hat{x}_k \quad if\ \ p_i(t) = 1 \quad (9)$$

where α^S is the update rate; $\hat{x}_k$ (k=1,2) is the human position estimated by a steady-state genetic algorithm using the 3D distance image sensor [20]. Furthermore, if the PSP of the estimated human position is less than the predefined threshold and if the temporal difference of human position is also small, then the position of the furniture or equipment nearest with the current position is used as the human location.

Figure 7 shows an experimental result of human extraction in the bed by the human detection in SSGA where (a) is the original illumination image, (b) is the original distance image, (c) and (d) are the plotting results of the side and front views, respectively, (e) is the reliability map for differential extraction, and (f) is the possible human area based on the differential extraction and the search result of SSGA for human detection depicted as a red color. The human position is calculated by the average of 3D positions of pixels corresponding to the red area.



(a) Illumination image (b) Distance image

(c) Side view of 3D space (d) Front view of 3D image

(e) Reliability map (f) Human search by SSGA

Fig.7. Image processing result of the extraction of the person in the bed [20]

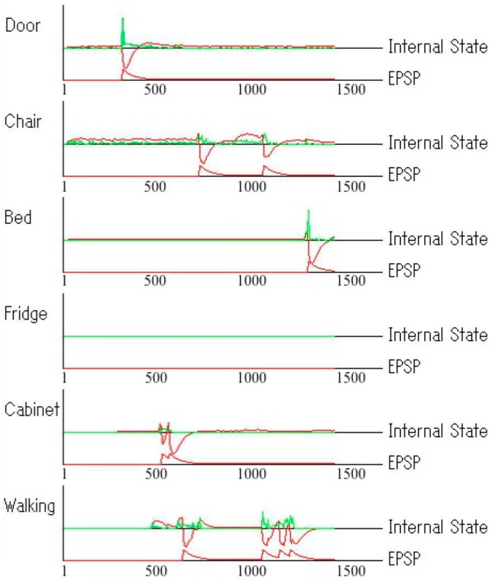


Fig.8. Presynaptic action potential and PSP over time

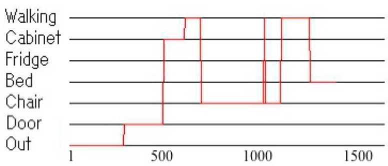
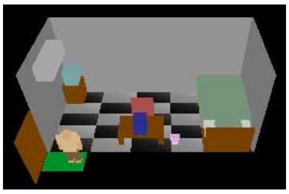
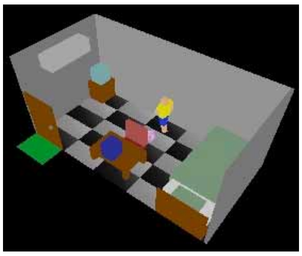


Fig.9. The transition of human position by SNN



(a) Entrance



(b) Walking

Fig.10. The visualization of human location on the iPhone and iPod touch in the remote observation system

3. EXPERIMENTAL RESULTS

This section shows several experimental results. We use the experimental example of a living room where the illumination sensors are attached with the fridge and cabinet, and the accelerometers are attached with the chair and bed shown in Fig.3. The number of spiking neurons is 6. These neurons are used for measuring the states of (1) entrance door, (2) chair, (3) bed, (4) cabinet, (5) fridge, and (6) walking. The human walking is extracted by the 3D distance image sensor [20].

Figure 8 shows experimental result of SNN where the green line is the input to each spiking neuron; the red line indicate the internal state (presynaptic action potential) of the upper figure, and the red line in the lower figure is the value of PSP in each neuron to measure the state of furniture or equipment. Each neuron fires as the state of internal state increases according to the sensory input calculated by the difference from the base value in each sensor. Figure 9 shows the estimated location of the person in the room. The person entered from the door, opened the cabinet, and took out a drink. After that, the person sat down the chair, and used the pocket robot partner to control home appliances. Finally, the person went to the bed (Figs.10 and 11 (4)).

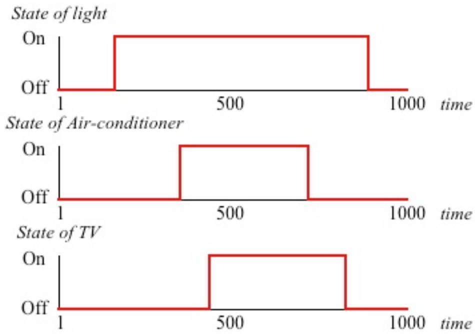
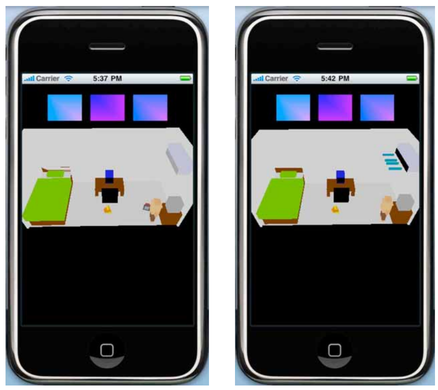


Fig.11. The state of home appliances



(a) Light (b) Air-conditioner

Fig.12. Operation results of universal remote controller based on pocket robot partner

Figure 10 shows the visualization of the human localization on the iPhone and iPod touch in the remote observation system. When the person goes home, the sensor in the entrance door fires. As a result, the person appeared in the entrance of the simulator (Fig.10 (a)). When the neuron corresponding to the walking fires, the person starts walking in the simulator according to the position detected by SSGA (Fig.10 (b)). Because the person is displayed as an abstract human model in the simulator, the privacy of the person can be protected. In this way, the sensing device of sensor networks and iPhone are connected through the informationally structured space.

Figure 11 shows the change of values on the state of home appliances. Figure 12 shows the operation result of universal remote controller using iPhone in the room. When the person enters in the room, the light is not turned on. After the person entered in the room, the person turned on the light and air-conditioner (Fig.11). As a result, the inside of the room is lightened up (Fig.12 (a)), and the wind is drawn from the air-conditioner (Fig.12 (b)). In this way, while the person can easily control home appliances, the family living far away from the person can understand the situation of the living room in the remote care.

4. SUMMARY

This paper discussed the applicability of informationally structured space to remote observation and remote control of home appliances. Next, we developed universal remote controller using iPhone based on GUI and CUI modes. Furthermore, we developed a human localization method based on sensor network. We applied a spiking neural network to extract the human position based on the sensor network. Next, we proposed the learning method of spiking neural network based on the time series of measured data. Experimental results show the effectiveness of the proposed methods. The developed system is available for the observation of human location in a living room. Furthermore, the experimental results show that the universal remote controller is useful and helpful for people to control home appliances.

As a future work, we intend to combine the voice recognition and visual perception of the robot partners to the total system to realize natural interaction and multi-modal communication with people. Furthermore, we will use other sensing devices to extract human state effectively.

References

[1] S.K.Card, T.P.Moran and A. Newell: The Psychology of Human-Computer Interaction, Lawrence Erlbaum Associates, 1983.

[2] Donald A. Norman, The Design of Everyday Things, Basic Books, 1988.

[3] S.Attygalle, M.Duff, T.Rikakis, Jiping.H, Low-cost, at-home assessment system with Wii Remote based motion capture, In Proceedings of the Conference on Virtual Rehabilitation, pp.168-174, 2008.

[4] N.Kubota, Y.Shimomura, Human-Friendly Networked Partner Robots toward Sophisticated Services for A Community, Proc. of SICE-ICCAS 2006, pp. 4861-4866, 2006.

[5] N.Kubota, K.Nishida, Cooperative Perceptual Systems for Partner Robots Based on Sensor Network, International Journal of Computer Science and Network Security, Vol.6, No.11, pp.19-28, 2006.

[6] N.Kubota, T.Obo, T.Fukuda, An Intelligent Monitoring System based on Emotional Model in Sensor Networks, The 18th IEEE International Symposium on Robot and Human Interactive Communication, pp.346-351, 2009.

[7] C. Corderio, D. Agrawal, Ad Hoc & Sensor Networks, World Scientific, 2006.

[8] I. Khemapech, I. Duncan and A. Miller, "A Survey of Wireless Sensor Networks Technology," Proc. of The 6th Annual PostGraduate Symposium on The Convergence of Telecommunications, Networking and Broadcasting, 2005.

[9] J-H. Kim, Y-D. Kim, and K-H. Lee, "The Third Generation of Robotics: Ubiquitous Robot," Proc. of 2nd International Conference on Autonomous Robots and Agents, 2004, pp.1-12.

[10] M.Satomi, H.Masuta, N.Kubota, Hierarchical Growing Neural Gas for Information Structured Space, IEEE Symposium Series on Computational Intelligence 2009.

[11] AW. Gerstner, Spiking Neurons, Pulsed Neural Networks, edited by W.Maass and C.M.Bishop, MIT Press, 1999, pp. 3-53.

[12] J.-S.R.Jang, C.-T.Sun, and E.Mizutani, "Neuro-Fuzzy and Soft Computing", Prentice-Hall Inc., 1997.

[13] J. A.Anderson, E.Rosenfeld, "Neurocomputing", The MIT Press, Cambridge, Massachusetts, 1988.

[14] N.Kubota, K.Nishida, Perceptual Control Based on Prediction for Natural Communication of A Partner Robot, IEEE Transactions on Industrial Electronics, Vol.54, No.2, pp.866-877, April, 2007.

[15] N.Kubota, K.Tomoda, "Behavior Coordination of A Partner Robot based on Imitation", Proc. of 2nd International Conference on Autonomous Robots and Agents, pp.164-169, 2004.

[16] N.Kubota, Computational Intelligence for Structured Learning of A Partner Robot Based on Imitation, Information Science, No. 171, pp.403-429, 2005.

[17] N.Kubota, K.Nishida, Development of Internal Models for Communication of A Partner Robot Based on Computational Intelligence, Proc. of 6th International Symposium on Advanced Intelligent Systems, 2005, pp. 577-582.

[18] N.Kubota, Y.Tomioka, M.Abe, Temporal coding in spiking neural network for gestures recognition of a partner robot, Proc. Joint 3rd Int. Conf. on Soft Computing and Intelligent Systems and 7th International Symposium on advanced Intelligent Systems, 2006, pp. 737-742.

[19] N.Kubota, A.Yorita, Topological Environment Reconstruction in Informationally Structured Space for Pocket Robot Partners, Proc. of 2009 IEEE International Symposium on Computational Intelligence in Robotics and Automation, 2009.

[20] N.Kubota, T.Obo, B.H.Lee, Localization of Human in Informationally Structured Space Based on Sensor Networks, Proc. of IEEE World Congress on Computational Intelligence 2010, 2010 (accepted).

[21] C.S.Chan, H.Liu, Fuzzy Qualitative Human Motion Analysis, IEEE Transactions on Fuzzy Systems, Vol.17, No.4, 851-862, 2009.

[22] H.Liu, A Fuzzy Qualitative Framework for Connecting Robot Qualitative and Quantitative Representations, IEEE Transactions on Fuzzy Systems, Vol.16, No.6, 1522-1530, 2008.