Robotica (2006) volume 24, pp. 643–648.
doi:10.1017/S0263574706002815 Printed in the United Kingdom

Multiagent cooperation based entertainment robot

Huang Yanwen†, Cao Qixin†,*, Zhou Jingliang†, Huang Yi† and Frank L. Lewis‡

†*Research Institute of Robotics, School of Mechanical Engineering, Shanghai Jiao Tong University, 1954 Huashan Road, Shanghai 200030 (P.R. China)*

‡*Automation and Robotics Research Institute, The University of Texas at Arlington, 7300 Jack Newell Blvd. S, Ft. Worth, Texas 76118 (USA)*

(Received in Final Form: March 2, 2006. First published online: May 4, 2006)

SUMMARY

In recent years, cooperation of heterogeneous multiagents (HMAs) has achieved formidable results and gained an increasing attention by researchers. This paper presents an entertainment robot system based on the HMAs' cooperation. This robot is used to distinguish the basic information of a person from the audience and entertain him in the exhibition hall. The main agents include the Arbitration System (AS), the Face Recognition System (FRS), the Position Cognition System (PCS), and the Behavior Planner (BP), which are carved up according to their respective functions. Each agent completes its own task independently and the robot finishes the whole mission through their cooperation. The hybrid control architecture and the data-fusion algorithm based on Bayesian Belief Network are also discussed in this paper.

KEYWORDS: Entertainment robot; Heterogeneous; Multiagent; Cooperation; Hybrid control architecture; Bayesian Belief Network; Data fusion.

1. INTRODUCTION

With the rapid growth of automation technology, robotics has involved into human lives from industry domain to daily life applications such as home-helpers or, more recently, entertainment robots. Entertainment robot, or E-robot is a kind of personal robot that is designed to co-exist with human beings in their everyday lives.[1,2] It is one of the most important emerging applications in the 21st century. It can be used at amusement arcades, exhibition halls, parties, sports events, or even homes. The market entry of AIBO[3] in 1999 is just the beginning of this new trend. Besides, the biped E-robots such as ASIMO,[4] SDR-3X,[5] and SDR-4X,[6] QRIO[7] may be typical examples of this kind of robots.

In future, E-robots will be more often applied to art, dance, film, and other types of performances such as plays than has been until now. The capability such as high performance face, speech recognition, and sophisticated behavior control for these E-robots will be more and more necessary. Considering the highly dynamic and uncertain environment that the E-robots have to interact with compared to industrial robots, designing the cognitive, motion, speech, behavior models and embedding them into one robot is not a simple design task as it includes machine learning, scene analysis, data fusion, context analysis, and high-level programming, etc.[7] It is also difficult to design one frame for different kinds of E-robots as they are specially designed for a specific application till now (e.g., AIBO and QRIO). The E-robots mentioned earlier have problems such as high cost, long design period, unique function, and the difficulties of reconfiguring. Building a cheap but functional E-robot will continue to be a major challenge. Considering the inherent advantages of multiagents system including the redundancy against the failure of the individual agent, the flexibility and resilience within real-world environments, the ability to specialize individual functions, and the engineering and development advantages, it is makes sense to construct the robot as an integrated system of multiagents, including sensors, execution, planning modules, etc.

We constructed a flexible robot system as a heterogeneous multiagent system (HMAS) which supports flexible connections among the independent agents and is easily expansible. Besides, the HMAS framework makes it possible to design a complex robot more easily and at low costs.

This paper presents the fundamental hardware and software system of the E-robot. The rest of the paper is organized as follows: Section 2 gives an overview of the robot's configurations. Section 3 discusses the robot's multiagents and their functions. The robot's control architecture and the cooperation between these multiagents are also discussed in this section. Section 4 discusses the data-fusion arithmetic. One example of the application is described in Section 5. Section 6 concludes this paper.

2. CONFIGURATION OF THE E-ROBOT

The E-robot is shown in Fig. 1. In this figure "1" denotes the robot's eyes composed of two CCD cameras used for face recognition; "2" is the robot's speaker; "3" is the LED used to show the robot's gender; "4" represents the robot's arm with three DOFs for simple action; "5" is the robot's knee for height adjustment; and "6" denotes two active and two passive wheels for its motion; "7" identifies two ultrasonic sensors for obstacle avoiding.

* Corresponding author.

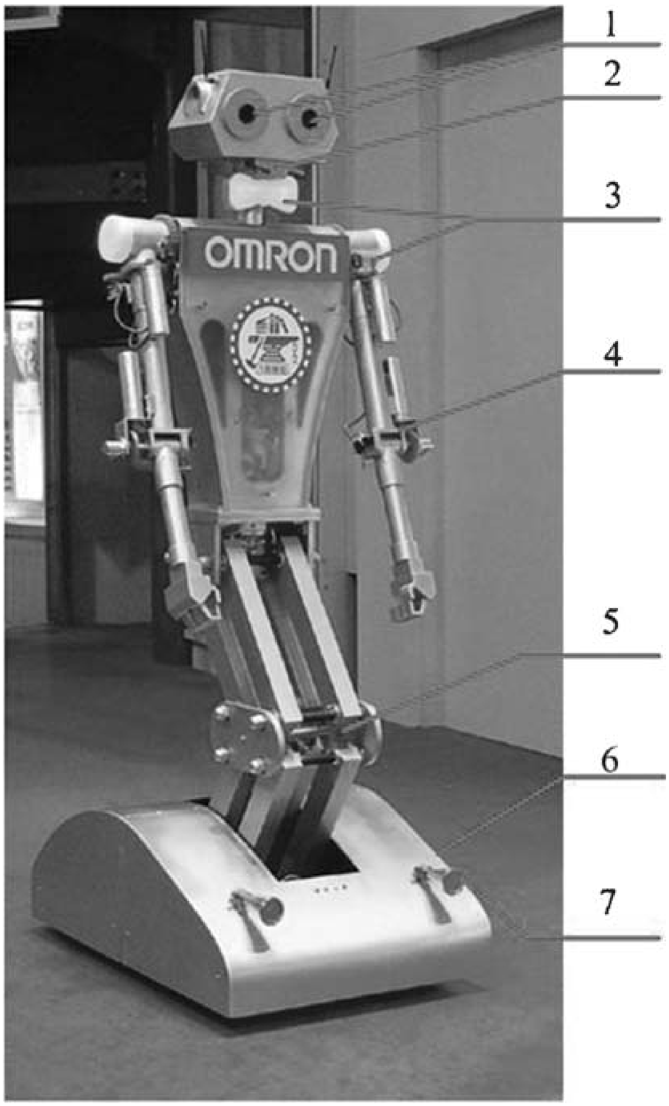

Fig. 1. E-robot's appearance and structure. 1 – robot's eyes composed of two CCD cameras used for face recognition, 2–robot's speaker, 3–LED used to show the robot's gender, 4–robot's arm with three DOFs for simple action, 5–robot's knee for height adjustment, 6–two active and two passive wheels for its motion, 7–two ultrasonic sensors for obstacle avoiding.

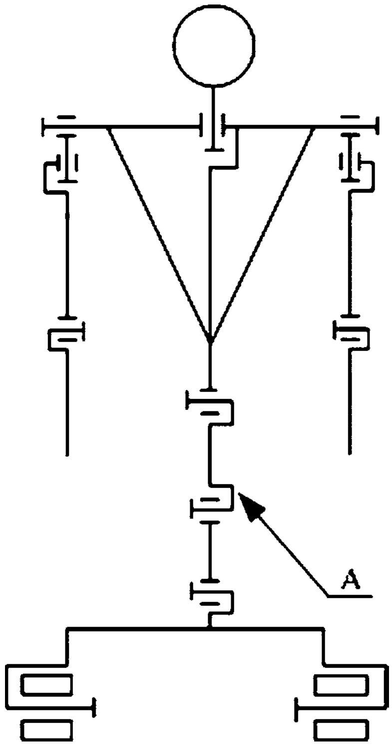

Fig. 2. Configuration of the robot's DOFs.

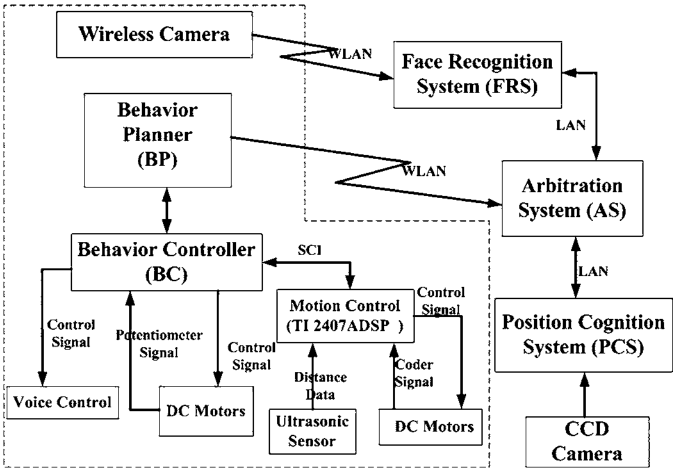

Fig. 3. Whole structure of the robot's control system.

The E-robot weighs approximately 100 kg, and its height ranges from 150 to 180 cm, which can be adjusted independently according to need. The degrees of freedom (DOFs) are 10 in total: neck (1), knee (1), arm (3×2: elbow, 1, shoulder, 2), and motion (2). The configuration of our E-robot's DOFs is shown in Fig. 2. All joints are active except the middle joint in the robot's knee, which is labeled as "A" in Fig. 2. The robot can regulate its speed from 0 to 1.5 m/s with two differential wheels. A total of 11 DC servo motors installed in the robot makes it possible to move smoothly as well as to behave freely. There are two cameras, two ultrasonic sensors, two coders, and nine potentiometer sensors fixed in the robot, which help it to perceive the environment and its own state.

The robot's control system is composed of the Arbitration System (AS), the Face Recognition System (FRS), the Position Cognition System (PCS), and the Behavior Planner (BP) according to their respective function. Each part communicates with other parts via LAN or Wireless LAN (WLAN). Each controller works as an intelligent agent, which completes its sub-task automatically and independently. The robot completes its whole mission through the agents' cooperation. The structure of the robot's control system is shown in Fig. 3.

3. HETEROGENEOUS MULTIAGENTS SYSTEM'S FUNCTIONS AND THEIR COOPERATION

3.1. Heterogeneous multiagents system's functions

In the real world, the E-robot often operates under real-time constraints that are inherently distributed and dynamic. Traditional static and hierarchical control structures do not suffice for such situations. A distributed real-time heterogeneous multiagent control system with the capabilities such as sensor data fusion, actuation processing, information exchange over a network, and synchronization to achieve the overall system goals can provide improved functionality, performance, flexibility, and reduced complexity and cost.

We divide our robot control system into several heterogeneous multi agents, which are the AS agent, the FRS agent, the PCS agent, and BP agent. The FRS agent is in charge of recognizing the gender and the age of the person facing the robot; the PCS agent is used for robot's self-localization and goal identification as well as obstacle avoidance; the BP agent is used to control the robot's behaviors and speech; and the AS agent is responsible for the sensor data fusion and the decision making. Each agent's function will be discussed in detail in the following sections.

3.1.1. The FRS agent. The human face is a rich source of information that the machines can use to understand people visually in the same way as humans do. The facial recognition is receiving increased attention due to its noticeable applications such as in security technologies.

The FRS agent, or the OKAO Vision System, which is developed by OMRON Corporation,[8] processes the acquired digital image to recognize the gender and age of the person facing the robot, based on the facial character information such as the eyes, the eyebrows, the nose, the lip, as well as the outline of the face. It then sends the results to the AS agent through WLAN. Then the robot automatically chooses its speech, appearance, and behavior based on such information. Further, the face image which is acquired by the FRS agent can help the robot to determine the direction and distance of the tracked objects such as the audience, to adjust its own position, pose, and height for better facial recognition effect and interaction between the robot and the human beings. The processing of face capture and face recognition and their results are shown in Fig. 4.

Fig. 4. Face recognition experiment. Processing of (a) face capture (b) face recognition.

Fig. 5. Localization of the robot by PCS.

3.1.2. The PCS agent. The PCS agent, which is put above the center of the stage, is used for robot self-localization and obstacle identification. The details of this algorithm are described in the following sections.

Firstly, the color study is run from the training images to extract the color blob of the goals in HSL color space, which provides better adaptability to the change of lightness. Secondly, after object recognition, the PCS agent picks up the circumrectangle, which embodies the interested color blob, which is the robot's logo in our application, in order to decrease computational complexity and to save the runtime. Thirdly, it processes the selected region through a binarization algorithm to distinguish the goal from the background. Then the centroid and the moment of inertia of the pixels (whose value is equal to 1) is computed. Finally, after computing the coordinate of the goal's center of mass, the robot's position is obtained. After computing the angle between the moment of inertia of the goal and the horizontal lines, the robot's pose is obtained. The processing result of the algorithm can be seen in Fig. 5. The (x, y, θ) of the robot is equal to (165, 286, 169).

3.1.3. The BP agent. The BP agent, which is equipped in the robot, controls the robot's behaviors and speech automatically according to the sensor data and the commands from the AS agent.

The BP agent is divided into a higher-level controller and a lower-level controller. We use an embedded PC104 as a higher-level controller for the behavior-based control and voice processing. An Inter-Process Communication (IPC) mechanism is introduced here to realize distributed software design, which is used to get sensor data, to arrange robot's behavior and to provide motor output commands and communication packets, as well as to process the voice. The TI DSP TMS320LF2407A is adopted here as a micro-controller, used for the motion control. Besides, the odometer data gained from coders is used for the robot's self-localization, and the distance data between the obstacle and itself measured by the ultrasonic sensors is used to avoid danger. The higher-level controller and the micro-controller are connected via RS-232 serial link. The BP agent communicates with the AS through WLAN that complies with the IEEE 802.11 standard. The architecture of the BP system is shown in Fig. 3.

3.1.4. The AS agent. The AS agent acts as the team coordinator to make decisions at the global task level, while each agent executes its own sub-task. In this way, the information transmitted among these agents can be minimized in order to meet the time demands of a highly dynamic environment. The AS agent receives sensor data from PCS, FRS, and the current robot state reported from BP through LAN or WLAN. It then processes and fuses the data, runs the whole process of the system, makes decisions, and finally sends commands to the BP agent.

3.2. The control algorithm of the multiagent system

We use the hybrid multilayer control architecture in our E-robot. The reactive control scheme is used in the BP agent as an executive layer and a behavior layer; the deliberative control scheme is used in the AS agent as a planner layer to determine the modalities of achieving higher-level tasks. Besides, the FRS and the PCS agents are independently employed as the perception and the monitor modules, respectively.

The reactive control strategy is devised here for the robot's behavior control. The BP agent takes data from sensors and gains commands from the AS agent, and then generates actions in response to these data directly. In order to realize a close connection between the perceived state of the world and the action chosen, no accurate model of the world is built or detailed response actions are defined. This has produced systems that are responsive and resilient, and cope well with more dynamic and unpredictable environments. However, this approach also has disadvantages: agents react only to local conditions and may not have the capability to achieve long-term goals efficiently. Therefore, we devised the AS agent for the global arbitration.

We use one rule-based algorithm to do the action selection in the AS agent, as the real world consists of numerous unexpected situations which cannot be finitely identified. In this method, stimulus is represented as a collection of "IF precondition THEN consequent". The precondition consists of a list of situations while the result is a behavior response. There are no predetermined situations that need to be recognized. This will ensure the feasibility of the system to be executed in the real world.

3.3. The multiagent's cooperation

In this section, we describe the design of multi heterogeneous agents, each one possessing different functions and discursion mechanisms. Agents are typically model-based, both in relation to their beliefs about the world in which they operate, and in relation to their knowledge of other agents. Each individual agent is able to carry out its tasks and to complete the system mission through interaction with other agents. In this way, the mission can be solved quickly by a number of agents operating in parallel. Besides, by duplicating the capabilities across multiagents, the robot has the potential to increase the robustness and reliability of the automated solution through cooperation. The data stream and the processing of the robot control system are shown in Fig. 6.

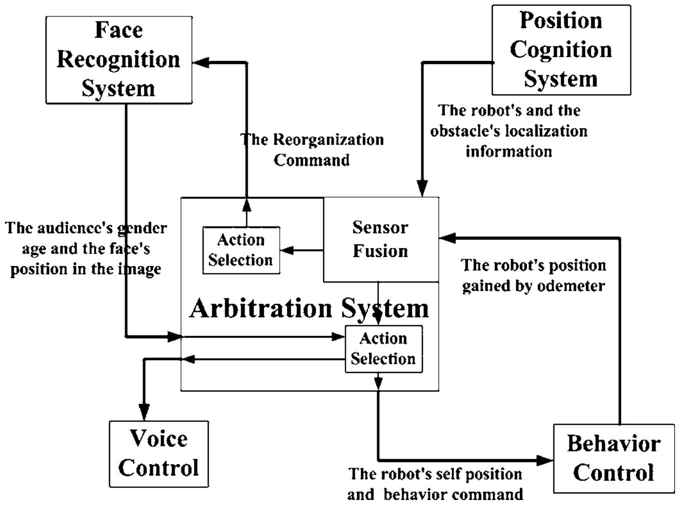

Fig. 6. Data stream and processing in the robot's control system.

4. MULTIDATA FUSION ALGORITHM

E-robots have to deal with unstructured, mutable, and generally unpredictable environments, and therefore require a higher degree of intelligence and flexibility. But constructing and updating models from one sensor to obtain accuracy and completeness is infeasible. Moreover, even without considering computational resource problems, the inherent limitation of sensor range puts a theoretical limit on the accuracy of the model that can be derived. Therefore, data fusion is often used to dynamically localize the robot and to construct representations of the dynamically changing environment.[9] Analyzing the uncertainties in the perception models of the vision system and the accumulative total fault in dead reckoning systems for the mobile robot,[10] a Bayesian Belief Network is proposed here to fuse data from odometer and the PCS for the robot's self-localization at successive observation points, as shown in Fig. 7.

In Fig. 7, $P_i(p_{xi}, p_{yi}, p_{\theta i})$ is the belief of the sensor's observations $R_i(x_i, y_i, \theta_i)$. Here, R_1 and R_2 denote the robot's pose observed by the global vision system and the odometer, respectively, R_3 denotes the obstacle's pose observed by the global vision system, R_4 denotes the obstacle's pose observed through the ultrasonic sensors, and P_i is decided on the ground of the information coming from the sensors, such as the vision system, the odometer, and ultrasonic sensors. $Q_1(x_1, y_1, \theta_1)$ is the global fused robot's

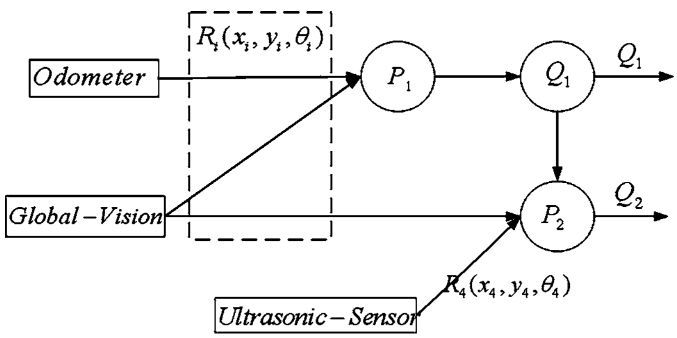

Fig. 7. Data fusion using Bayesian belief network.

Fig. 8. Scene of E-robot's performance.

position and $Q_2(x_2, y_2, \theta_2)$ is the global fused obstacle's position.

The first stage is the robot's global pose fusion as shown in Fig. 7. The belief of the vision sensor's observations and the coder sensor's observations are calculated, respectively. Then, the global fused robot position and the belief are calculated, which is denoted as Q_1 in Fig. 7.

Once we get the global fused robot position, we can use it to calculate the obstacle's position Q_2 together with the data observed by the vision system and the ultrasonic sensor and their corresponding beliefs.

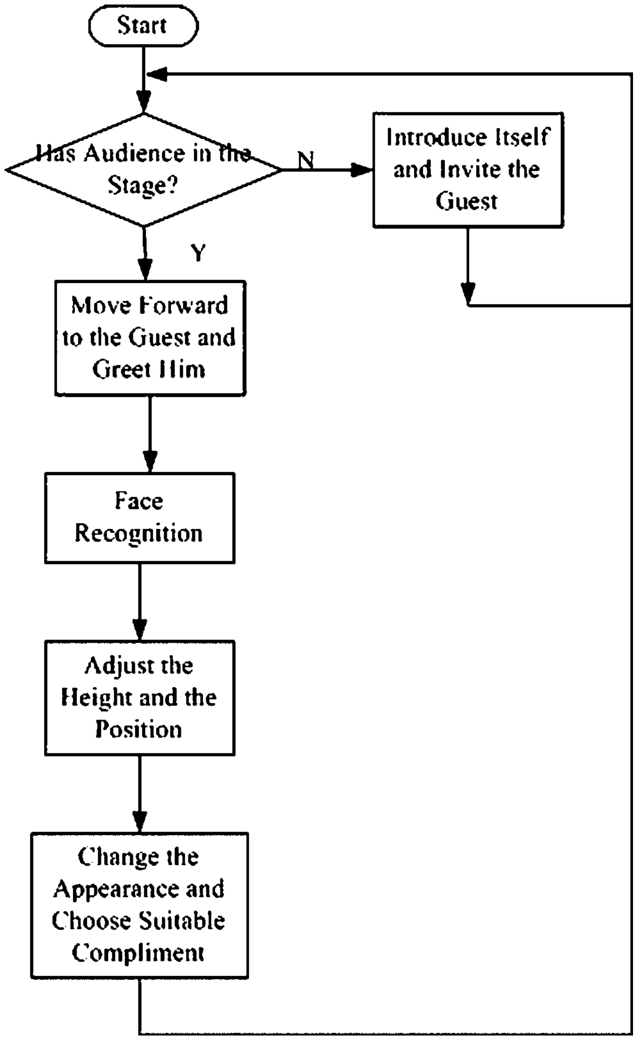

Fig. 9. Flow chart of the robot's performance.

5. EXHIBITION OF THE E-ROBOT

This robot has been created to communicate with and to entertain visitors in the exhibition.

In the exhibition hall, the robot oriented its position firstly by fusing the sensor data in the AS agent gained from PCS agent and the odometer data transmitted from BP agent. When a person from the audience was invited on the stage, the PCS agent recognized his position, and then sent this information to the BP agent through WLAN to control the robot's movement towards the person to welcome him. Two ultrasonic sensors positioned on the front of the robot could measure the distance between the obstacle and the robot in order to help the BP agent to localize the visitor and to avoid danger. After getting images from the FRS agent through WLAN, the robot imperceptibly adjusted its position, pose and height for better facial recognition effect. Then, according to the information including the audience's gender and age transmitted from the FRS agent, the robot changed the color of LED and said different words to the visitor, such as "Nice to meet you, beautiful lady", or "Welcome gentleman, have a good time". The scene and the flow of the robot's performance are shown in Figs. 8 and 9, respectively.

6. CONCLUSIONS

We have presented an E-robot system based on heterogeneous multiagents' cooperation. In particular, we have described the robot's hardware and software configuration and provided an example of application. As the E-robot is a symbolic product that has already established a milestone for the new wave of robotics building in the new century, its required functions may vary according to the particular application. The multiagent configuration makes it easy to accede or detract one function to the whole robot system (such as adding the voice recognition, or increasing the DOFs) without redesigning. The cooperation among those multiagents makes the robot's design much cheaper and more practical in many applications since building a certain number of less capable agents that can work together to accomplish a mission is much more reliable and cheaper than trying to build one complex agent in charge of performing the entire mission.

Acknowledgments

Support for this research was provided in part by the National Natural Science Foundation of China under Grant

No. 60443007. The development of the robot was carried out in collaboration with OMRON Corporation.

References

1. T. Makimoto and T. T. Doi, “Chip Technologies for Entertainment Robots – Present and Future”, *Technical Digest – International Electron Devices Meeting* (2002) pp. 9–16.
2. M. Asada, R. D’Andrea, A. Birk, H. Kitano and M. Veloso, “Robotics in Edutainment”, *Proceedings of the IEEE International Conference on Robotics and Automation*, San Francisco, California, USA (April 2000) **1**, pp. 795–800.
3. R. C. Arkin, M. Fujita, T. Takagi and R. Hasegawa, “Ethological modeling and architecture for an entertainment robot,” *Proceedings of the IEEE International Conference on Robotics and Automation*, Seoul, Korea (2001) **1**, pp. 453–458.
4. K. Hirai, M. Hirose, Y. Haikawa and T. Takenaka, “The development of Honda humanoid robot,” *Proceedings of the IEEE International Conference on Robotics and Automation*, Leuven, Belgium (1998) **2**, pp. 1321–1326.
5. Y. Kuroki, “A small biped entertainment robot,” *Proceedings of the 2001 International Symposium on Micromechatronics and Human Science,* MHS 2001, Nagoya (Sept. 9–12, 2001) pp. 3–4.
6. T. Ishida, Y. Kuroki and J. Yamaguchi, “Mechanical system of a small biped entertainment robot,” *Proceedings of the IEEE International Conference on Intelligent Robots and Systems*, Las Vegas, NV, USA (IROS 2003) **2**, pp. 1129–1134.
7. K. Nagasaka, Y. Kuroki, S. Suzuki, Y. Itoh and J. Yamaguchi. “Integrated motion control for walking, jumping and running on a small bipedal entertainment robot,” *Proceedings of the 2004 IEEE International Conference on Robotics and Automation*, New Orleans, LA, USA (ICRA 2004) **4**, pp. 3189–3194.
8. http://www.omron.com/r_d/vision/01.html
9. A. T. Tokuhiro and B. J. Vaughn, “Initial test results of the Omron Face Cue Entry System at the University of Missouri-Rolla Reactor,” *J. Nucl. Sci. Technol.* **41**(4), 502–510 (2004).
10. A. W. Stroupe, M. C. Martin and T. Balch, “Distributed sensor fusion for object position estimation by multirobot systems,” *Proceedings of the IEEE International Conference on Robotics and Automation*, Seoul, Korea (2001) **2**, pp. 1092–1098.