Proceedings of the 2010 IEEE
International Conference on Robotics and Biomimetics
December 14-18, 2010, Tianjin, China

Human-Robot Interaction by Reading Human Intention based on Mirror-Neuron System

Ji-Hyeong Han and Jong-Hwan Kim

***Abstract*— Considering the human-robot symbiosis in coming years, robots should be able to read human intention for natural and rational interaction with human beings. This paper proposes effective human-robot interaction (HRI) by reading the human intention using cognitive architecture. Proposed intention reading algorithm is inspired by mirror-neuron system and simulation theory which are the significant parts of human mind reading skill. For human intention reading, the cognitive architecture consisting of eight modules, i.e. perception, attention, behavior mapping, behavior model, intention reading, memory, behavior selection, and actuator modules is also proposed. The effectiveness of the proposed scheme is demonstrated through computer simulations on human-robot play with two different objects, such as a ball and a toy car.**

I. Introduction

These days it is expected that the era of human-robot symbiosis is rapidly approaching along with the rapid improvement of robot technology and artificial intelligence (AI). To keep the relationship between human and robot that live close together and depend on each other in particular ways in a certain environment, natural and rational human-robot interaction (HRI) is needed. Through the effective HRI, each can get particular benefits from the other. Thus, the research on HRI is required and it has become one of the fast growing fields in robotics.

Reading human mind and intention is a significant part of HRI research for effective interactions between human and robot. The robot can recognize the human intention by understanding human's verbal command. In this case, however, human has to make an explicit command to the robot all the time. To solve this problem and to have natural HRI, the robot has to have a capability to read the human intention by observing his/her behavior. For this purpose, there were several researches on intention reading algorithm. Schrempf et al. proposed a generic model for estimating user intentions using hybrid dynamic bayesian networks [1], [2]. Schmid et al. proposed proactive robot task selection given a human intention estimate [3]. Omori et al. proposed a light loaded computational algorithm that achieved human-robot interaction without intention estimation in the self agent, but assuming the other agent to estimate intention [4]. Gray et al. and Breazeal et al. proposed action parsing and goal inference algorithm using self as simulator and an embodied cognition of mind-reading skills [5], [6].

The authors are with the Department of Electrical Engineering, KAIST, 373-1 Guseong-dong, Yuseong-gu, Daejeon 305-701, Republic of Korea (e-mail: {jhhan, johkim}@rit.kaist.ac.kr).

The ideal way of HRI is imitating the interaction between humans, since the action and reaction of the robot need to be natural and rational to humans. Humans interact with others very effectively using various ways such as language, gesture, facial expression, etc. Gesture and non-verbal cues are the important parts of effective interaction. Since humans understand the meaning of other's gesture or non-verbal cues, they can read his/her mind including intention and goal and even predict his/her next one. The major causes of mind-reading skill of humans discovered from neuroscience are mirror-neuron system and simulation theory (ST) [7]. Mirror-neuron system is activated both when individual performs a particular action and when individual observes the same action performed by others. ST is one of mind-reading theories, which is closely related with the mirror-neuron system. It suggests that humans use their own mental mechanism to read other's mind like simulation.

This paper proposes a human intention reading algorithm which is inspired by mirror-neuron system and ST. The basic concept is the reading of the human intention by observing his/her action and simulating it on robot's own actuator system. However, there is a case that the human intention is ambiguous by observing current action only. To deal with this case, uncertainty level (UL), trial behavior, and human feedback are introduced similarly as in human-human interaction. For intention reading, the cognitive architecture consisting of eight modules, i.e. perception, attention, behavior mapping, behavior model, intention reading, memory, behavior selection, and actuator modules is provided. The attention module makes the robot pay attention to moving objects including humans using interest factor (IF). The behavior mapping module maps a human behavior to robot's own behavior model and makes the robot recognize the human behavior. The intention reading module is to read the human intention using the proposed human intention reading algorithm. The memory module makes the robot save the information of selected behavior and reuse it later. To demonstrate the effectiveness of the proposed intention reading algorithm and cognitive architecture, computer simulations are carried out for human-robot interaction when they are playing with a ball and a toy car.

This paper is organized as follows. Section II briefly describes the mirror-neuron system and two major theories of mind-reading. Section III proposes the human intention reading algorithm and the overall cognitive architecture. In Section IV, a simulation environment is presented and simulation results are discussed. Finally, concluding remarks follow in Section V.

978-1-4244-9318-0/10/$26.00

II. Human's mirror-neuron system for mind-reading

Humans can interact with others without using explicit words. Since humans can read other's mind, such as intention, goal, belief, etc., gesture and facial expression are very important parts of interacting with understanding other people. There is a lot of evidence that mirror-neuron system is a part of mind-reading process. Also, there are two major theories of mind-reading, i.e. theory theory (TT) and simulation theory (ST). ST is closely related with the mirror-neuron system. The following briefly descries the mirror-neuron system and the two theories.

A. Human Mirror-Neuron System

Mirror neuron, which was originally discovered in area F5 of the monkey's premotor cortex, is a class of visuomotor neurons [7]-[9]. These neurons respond both when a particular action is performed by itself and when the same action, performed by another individual, is observed. There is a lot of indirect evidence that the mirror-neuron system is also in humans from neurophysiological and brain-imaging experiments. Neurophysiological experiments showed that when humans observed an action done by another their motor cortex became active without any motor activity [10]-[12]. Brain imaging studies demonstrated that the observation of actions done by others activated human's complex network formed by occipital, temporal, and parietal visual areas and two cortical regions whose function was fundamentally or predominantly motor [13]-[15]. There are two possible functions of the mirror system: one is learning by imitation and the other one is mind-reading or precursor to such a process. Mind-reading means representing mental states of others such as intention, goals, beliefs, etc. Identifying other's intention is useful for effective interaction with them, since it helps him/her anticipate their next action.

B. Two Theories of Mind-Reading

There are two dominant theories of mind-reading: theory theory (TT) and simulation theory (ST) [16], [17]. The fundamental idea of TT is that humans complete mind-reading by obtaining and organizing a commonsense theory of mind. According to TT, mental states to others arise from theoretical reasoning involving known causal laws. On the other hand, ST suggests that humans use their own mental mechanism to predict the mental processes of others like simulation. According to ST, humans observe other's action and simulate for backward inference from the observed action to a goal state. ST is in a similar fashion with mirror-neuron system and does not need any prior knowledge, database or common causal laws for mind-reading.

III. Cognitive architecture for intention reading

The intention reading is a necessary part of cognition for natural and effective HRI. In this section, an intention reading algorithm which is inspired by mirror-neuron system and simulation theory (ST) is proposed. For intention reading, cognitive architecture is also provided.

A. Cognitive Architecture

Fig. 1 shows the overall proposed cognitive architecture, which consists of eight modules, i.e. perception, attention, behavior mapping, behavior model, intention reading, memory, behavior selection, and actuator modules. The information on objects and a human behavior are forwarded to the perception module from environment and the percepts are passed to the attention module. The robot pays attention to moving objects including humans by interest factor (IF) which is explained in the following subsection. Then the object information under attention is saved in the object submodule of memory module and the information of the human behavior is transferred to the behavior mapping module. Behavior mapping module maps the received behavior information to robot's own behavior model and finds a most similar behavior. Now the robot recognizes the human behavior based on its own behavior model. This mapped information is transferred to the intention reading module. In intention reading module, the human intention is recognized using the proposed human intention reading algorithm which is inspired by mirror-neuron system and simulation theory (ST). More details of the algorithm are explained in the following subsection. When human's final intention is identified in intention reading module, behavior selection module selects a most proper behavior considering his/her intention for natural HRI. The selected behavior is activated through actuator module. Each and every mapped and selected behavior in intention reading module is stored in the behavior submodule of memory module, which is linked with the object submodule. Fig. 2 shows the structure of memory module.

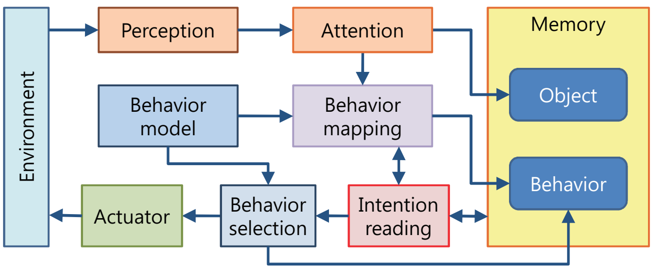


Fig. 1. The overall cognitive architecture.

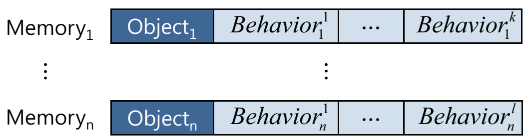


Fig. 2. The structure of memory module. n is the number of objects which are saved in the object submodule of memory module and k and l are the numbers of behavior which are saved in the behavior submodule of memory module linked with the object.

B. Interest Factor and Uncertainty Level

The robot needs to give the objects and humans in environment its attention for interacting with them. For this purpose, interest factor (IF) is defined as follows:

$$IF = \begin{cases} 1 & when\ object\ is\ moving \\ c_1 + c_2 e^{-k_1/\tau} & when\ object\ is\ not\ moving \end{cases} \tag{1}$$

where τ is a time constant, k_1 increases by one time step from the moment when object is not moving, and c_1 and c_2 are positive constants. When an object moves, IF for it becomes one immediately and the robot gives attention to the one among others which has the highest IF. When the object or human stop, IF is decreased by τ.

When the robot is in lack of information for recognizing human intention by observing current action, it needs to request more information to humans. For this purpose, uncertainty level (UL) is defined as follows:

$$UL = \begin{cases} 0 & when\ no\ more\ information\ is\ needed \\ c_3 \times k_2^2 & when\ more\ information\ is\ needed \\ 1 & when\ UL \geq 1 \end{cases} \tag{2}$$

where k_2 increases by one time step from the moment when the robot needs more information and c_3 is a positive constant. When the robot does not need more information, UL is zero. On the other hand, when it needs more information, UL increases proportional to the square of time step.

C. Intention Reading Algorithm

Fig. 3 shows the flowchart of the proposed intention reading algorithm. When the mapped information is transferred from behavior mapping module to intention reading module, if memory module does not have the information about the object under attention the algorithm starts to search higher behavior (HB) using robot's own behavior model like mirror neuron system and simulation theory. If the object information is already in memory module, the algorithm starts to search the memory for HB which is already linked with the object.

There are three cases when memory module does not have the information about the object under robot's attention. The first case is that the number of searched HB, $n(HB)$ is one. In this case, the algorithm repeats searching for HB of the current searched behavior. The second case is that $n(HB)$ is more than one and uncertainty level (UL) is zero. In the second case, since there are more than one searched HBs, the robot can not determine one of them by itself. Thus, the robot does a trial behavior for each searched HB and receives human feedback. The searched HB with the higher human feedback is determined as a HB and the algorithm repeats searching for HB of the determined one. The last case is $n(HB)$ is more than one and UL is bigger than the predefined threshold. In this case, the robot needs more information to search for HB, since UL is bigger than the predefined threshold. For this purpose, it requests a person another behavior. After his/her behavior, the algorithm goes back to behavior mapping module and repeats the process.

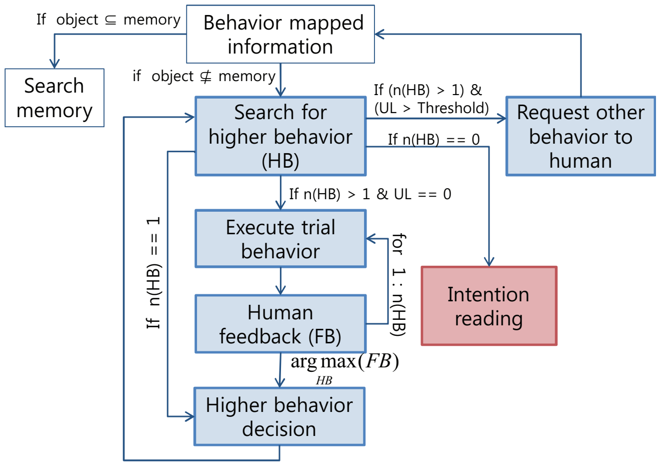


Fig. 3. The intention reading algorithm. HB is higher behavior, FB is feedback, UL is uncertainty level, and $n(x)$ is the number of x.

Finally, when there is no more HB, the intention reading algorithm is terminated and the last HB is determined as human intention. This whole process mimics the human intention reading process such as mirror-neuron system and simulation theory.

IV. Simulation

To demonstrate the effectiveness of the proposed intention reading algorithm along with cognitive architecture, human-robot play with two objects – a ball and a toy car – was employed as an application example. The robot needs to know the intention of human's behavior without explicit verbal command, like "going to the ball" or "throwing the ball," for natural and efficient human-robot play and interaction. There are three cases in simulation scenario: Case 1 is human-robot play with a ball, Case 2 is human-robot play with a toy car, and Case 3 is human-robot play with both the ball and the toy car. The following describes the simulation environment and simulation results.

A. Simulation Environment

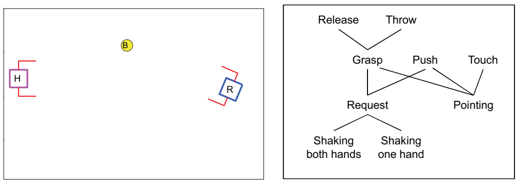
Human-robot play with two objects – a ball and a toy car – was simulated in 2D plane by MATLAB. Fig. 4(a) shows the simulation environment with a ball for Case 1. The robot was assumed to have a vision system and two manipulators. The hierarchical behavior model of the robot is described in Fig. 4(b). The behavior which is at upper position is a higher behavior (HB) of the one which is placed in a lower position and linked to it. The constants of IF and UL are defined as $c_1 = 0$, $c_2 = 1$, $c_3 = 1/225$, $\tau = 10$, and $threshold = 0.7$.

B. Simulation Results

1) Case 1: Fig. 5 shows the snapshots of simulation in chronological order.

i), ii) A person threw a ball.

iii) The robot perceived both of the person and the ball and approached the ball, since the ball was moving and IF of the ball was increasing. The information on the ball was



(a) (b)

Fig. 4. (a)Simulation environment. 'H' is a human, 'R' is a robot, and 'B' is a ball. (b)The hierarchy behavior model of the robot.

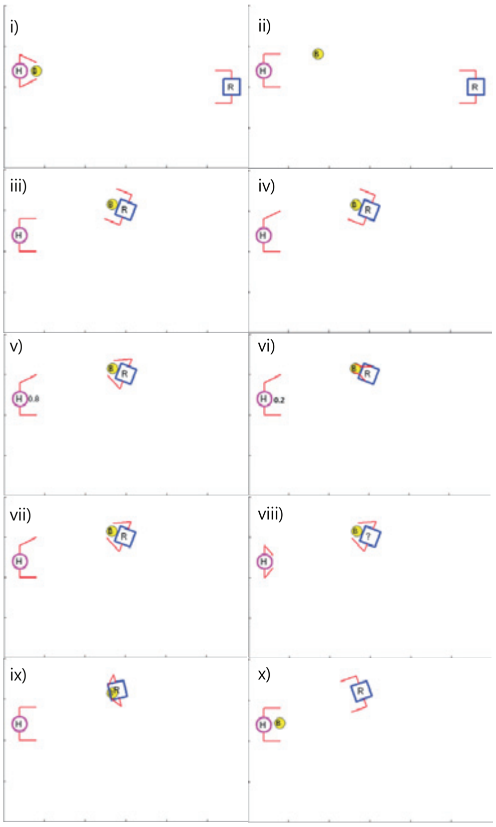


Fig. 5. The snapshots of simulation results of Case 1.

saved in the object submodule of memory module.

iv) When the robot arrived at the ball, the person was shaking one hand. The robot mapped the human behavior to its own behavior model in behavior mapping module and then recognized it was "shaking one hand." This information was transferred to intention reading module and the robot started to find the human intention using the proposed intention reading algorithm.

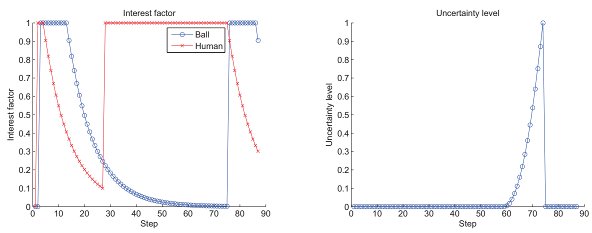
v), vi) The robot searched HB for "shaking one hand." Since "request" was the only valid HB, it was decided as HB. There were two possible HBs of "request," i.e. "grasp" and "push." The robot did trial behavior of "grasp" and was waiting for human feedback. The human feedback of "grasp" was 0.8. After the human feedback, the robot did another trial behavior of "push" and was waiting for human feedback. The human feedback of "push" was 0.2.

vii) "Grasp" which had higher human feedback was decided as HB.

viii) There were two possible HBs of "grasp" again, but the person still did the same behavior "shaking one hand". Therefore, the robot needed more information and UL was increasing. When UL was bigger than a threshold value, the robot asked the person more information by showing a question mark. The person did throwing behavior. Since the person did a new behavior and the robot did not need more information, UL was decreasing. The robot mapped the new behavior again and then recognized it was "throw."

ix), x) "Throw" was decided as HB. Since there was no more searched HB, "throw" was the final HB and it was the human intention. Thus, the robot threw the ball to the person.

Through this process using the proposed algorithm, the robot could play with a person with a ball without any preliminary knowledge of playing with a ball and explicit commands. Fig. 6(a) and Fig. 6(b) show the change of IF and UL, respectively. Fig. 7 shows the behavior which was decided as HB. These identified HBs were saved in the behavior submodule of memory module for later use. Therefore, the information on the ball, which was already saved in the object submodule of memory module at step iii), and the decided HBs were saved and linked together in memory module.



(a) (b)

Fig. 6. Case 1. (a)The interest factor. (b)The uncertainty level.

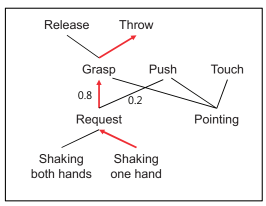


Fig. 7. The decided HBs of Case 1.

2) Case 2: Fig. 8 shows the snapshots of simulation in chronological order.

i), ii) A person pushed a toy car.

iii) The robot perceived both of the person and the toy car and approached the toy car, since the toy car was moving

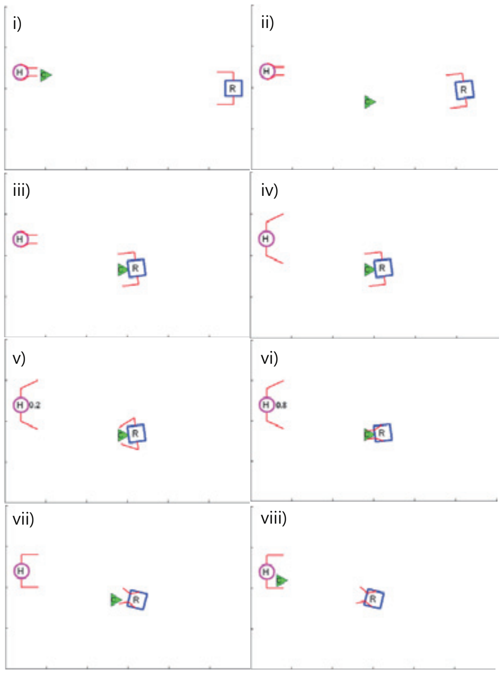


Fig. 8. The snapshots of simulation results of Case 2.

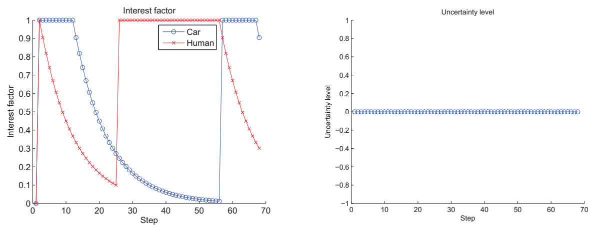
and IF of the toy car was increasing. The information on the toy car was saved in the object submodule of memory module.

iv) When the robot arrived at the toy car, the person was shaking both hands. The robot mapped the human behavior to its own behavior model in behavior mapping module and then recognized it was "shaking both hands." This mapped information was sent to intention reading module and the robot started to find the human intention using the proposed intention reading algorithm.

v), vi) The robot searched HB for "shaking both hands" and then found one valid HB, "request." Thus, "request" was decided as HB. Since there were two possible HBs of "request," i.e. "grasp" and "push," the robot did trial behavior of "grasp" firstly and was waiting for human feedback. The human feedback of "grasp" was 0.2. After the human feedback, the robot did another trial behavior of "push" and was waiting for human feedback. The human feedback of "push" was 0.8.

vii), viii) "Push" which had higher human feedback was decided as HB. Since there was no more searched HB, "push" was the final HB and it was considered as the human intention. Thus, the robot pushed the toy car to the person.

Through this process using the proposed algorithm, the robot could play with a person with a toy car and a person without any preliminary knowledge of playing with a toy car and explicit commands. Fig. 9(a) and Fig. 9(b) show the change of IF and UL, respectively. Since there was no moment when the robot needed more information, the value of UL did not change and remained zero. Fig. 10 shows the behavior which was decided as HB. These identified HBs were saved in the behavior submodule of memory module for later use. Therefore, the information on the toy car, which was already saved in the object submodule of memory module at step iii), and the decided HBs were saved and linked together in memory module.



(a) (b)

Fig. 9. Case 2. (a)The interest factor. (b)The uncertainty level.

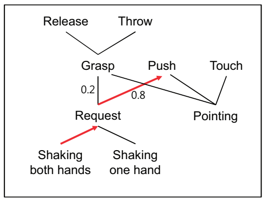


Fig. 10. The decided HBs of Case 2.

3) Case 3: Fig. 11 shows the snapshots of simulation in chronological order.

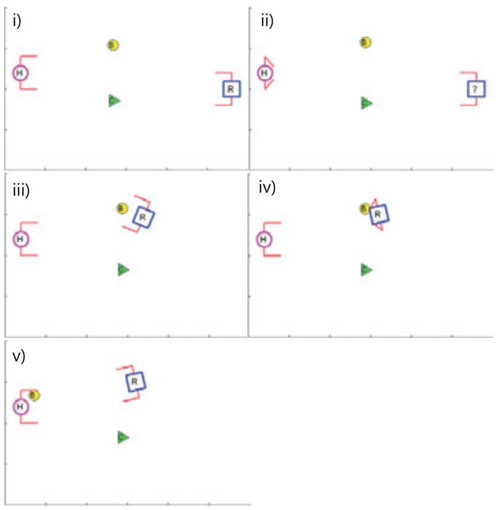


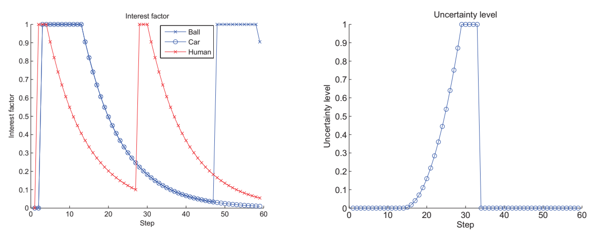
Fig. 11. The snapshots of simulation results of Case 3.

i) A person threw the ball and also pushed the toy car.

ii) The robot perceived the person, the ball, and the toy car simultaneously. The IFs of both the ball and the toy car were the same, since both of them were moving. Therefore, the robot could not decide where it would approach and UL was increasing because it needed more information about the moving objects. When UL was bigger than a threshold value, it asked the person more information by showing a question mark. Then, the person did throwing behavior.

iii), iv), v) Since the person did a new behavior and the robot did not need more information, UL was decreasing. The robot mapped the new behavior and recognized it was "throw." The robot searched memory module, since the ball and the toy car were already saved in the object submodule of memory module. Then, the robot found "throw" behavior was linked with the ball object. Therefore, the robot recognized that the human intention was throwing the ball. Finally, the robot approached the ball and threw it to the person.

Through this process using memory module, the robot could recognize the human intention and interact with humans naturally without asking an explicit verbal command in a confusing situation like that with two objects. Fig. 12(a) and Fig. 12(b) show the change of IF and UL, respectively. Fig. 13 shows the behavior which was decided as HB.



(a) (b)

Fig. 12. Case 3. (a)The interest factor. (b)The uncertainty level.

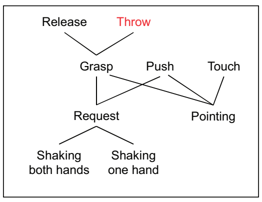


Fig. 13. The decided HBs of Case 3.

V. Conclusion

In this paper, the effective HRI by reading the human intention based on cognitive architecture was proposed. The proposed intention reading algorithm mimicked human's mind reading skill. It was inspired by mirror-neuron system and simulation theory which are important parts of human's mind reading. For human intention reading, the cognitive architecture consisting of eight modules was also proposed. The proposed scheme identified the human intention by observing a human behavior and simulating it on the robot's own behavior model. To deal with ambiguous cases, uncertainty level, trial behavior, and human feedback were also introduced. The effectiveness of the proposed scheme was demonstrated through computer simulations carried out for human-robot play with two different objects, i.e. the ball and the toy car. The robot could identify human intention and select proper behaviors for natural and effective HRI without explicit commands and preliminary knowledge about humans and objects.

Acknowledgment

This work was supported (National Robotics Research Center for Robot Intelligence Technology, KAIST) by Ministry of Knowledge Economy under Human Resources Development Program for Convergence Robot Specialists.

References

[1] O. C. Schrempf, U. D. Hanebeck, A. J. Schmid, and H. Wörn, "A novel approach to proactive human-robot cooperation," in *Proc. of the 14th IEEE Ro-Man05*, 2005.

[2] O. C. Schrempf and U. D. Hanebeck, "A general model for estimating user intentions in human-robot cooperation," in *Proc. of the 2nd International Conference on Informatics in Control, Automation and Robotics*, 2005.

[3] A. J. Schmid, O. Weede, and H. Wörn, "Proactive robot task selection given a human intention estimate," in *Work shop on Robot and Human Interactive Communication*, 2007.

[4] T. Omori, A. Yokoyama, H. Okada, S. Ishikawa, and Y. Nagata, "Computational modeling of human-robot interaction based on active intention estimation," *Lecture Notes in Computation Science*, vol. 4985, pp. 185-192, 2008.

[5] J. Gray, C. Breazeal, M. Berlin, A. Brooks, and J. Lieberman, "Action parsing and goal inference using self as simulator," *Proc. of the 14th IEEE Ro-Man05*, 2005, pp. 202-209.

[6] C. Breazeal, J. Gray, and M. Berlin, "An embodied cognition approach to mindreading skills for socially intelligent robots," *Internationl Journal of Robotics Research*, vol. 28, no. 5, pp. 656-680, 2009.

[7] V. Gallese and A. Goldman, "Mirror neurons and the simulation theory of mind-reading," *Trends in Cognitive Sciences*, vol. 2, no. 12, pp. 493-501, 1998.

[8] G. Rizzolatti and L. Craighero, "The Mirror-Neuron System," *Annual Review of Neuroscience*, vol. 27, pp. 169-192, 2004.

[9] V. Gallese, "Motor abstraction: a neuroscientific account of how action goals and intentions are mapped and understood," *Psychological Research*, vol. 73, no. 4, pp. 486-498, 2009.

[10] L. Fadiga, L. Fogassi, G. Pavesi, and G. Rizzolatti, "Motor facilitation during action observation: a magnetic stimulation study," *Journal of Neurophysiology*, vol. 73, pp. 2608-2611, 1995.

[11] F. Maeda, G. Kleiner-Fisman, and A. Pascual-Leone, "Motor facilitation while observing hand actions: specificity of the effect and role of observer's orientation," *Journal of Neurophysiology*, vol. 87, pp. 1329-1335, 2002.

[12] S. Patuzzo, A. Fiaschi, and P. Manganotti, "Modulation of motor cortex excitability in the left hemisphere during action observation: a single and paired-pulse transcranial magnetic stimulation study of self- and non-self action observation," *Neuropsychologia*, vol. 41, pp. 1272-1278, 2003.

[13] G. Buccino, F. binkofski, G. R. Fink, L. Fadiga, L. Fogassi, et al., "action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study," *European Journal of Neuroscience*, vol. 13, no. 2, pp. 400-404(5), 2001.

[14] J. Decety, T. Chaminade, J. Grezes, and A. N. Meltzoff, "A PET exploration of the neural mechanisms involved in reciprocal imitation," *NeuroImage*, vol. 15, pp. 265-272, 2002.

[15] J. Grèzes, J. L. Armony, J. Rowe, and R. E. Pasingham, "Activations related to "mirror" and "canonical" neurones in the human brain: an fMRI study," *NeuroImage*, vol. 18, pp. 928-937, 2003.

[16] M. Davies and T. Stone, "Folk psychology and mental simulation," *Royal Institute of Philosophy Supplement*, vol. 43, pp. 53-82, 1998.

[17] P. Carruthers and P. K. Smith, "Theories of theories of mind," *Cambridge University Press*, 1996.