2005 IEEE/RSJ International Conference on Intelligent Robots and Systems

Human-Robot Cooperative Sweeping by Extending Commands Embedded in Actions

Kazuki KOBAYASHI
Department of Informatics
The Graduate University for Advanced Studies (Sokendai)
2-1-2 Hitotsubashi, Chiyoda, Tokyo, Japan
kazuki@grad.nii.ac.jp

Seiji YAMADA
National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda, Tokyo, Japan
seiji@nii.ac.jp

***Abstract*— In this paper, we propose a novel interaction model for a human-robot cooperative task. The CEA (Commands Embedded in Actions) model reduces a human work-load because a user of a robot needs less inputs and outputs than DCM (Direct Commanding methods) like gesturing. We propose ECEA (Extended CEA) in order to deal with more complicated tasks than CEA. On the cooperative sweeping task between a human and a mobile robot, we apply temporal extension as one of ECEA instances. Mmultiple commands are embedded in the human action by the extention and the robot performs more complicated task. The experiments for conforming reduction of a human work-load using ECEA are conduced on the sweeping task. Human cognitive loads are measured as human work-loads and compared between ECEA and DCM. The results of the experiments showed that the ECEA minimized a human cognitive load.**

***Index Terms*— interaction design, cooperative task, sweeping, cognitive load, mobile robot**

I. Introduction

There are robots spreading among people as a progression of its technologies. We can purchase pet robots like AIBO [1] or cleaning robots like Roomba, and interact with them in a home environment. We will see tour-guide robots [2] in a museum in the near future. Robots thus have transferred their scene from industrial environments to home environments. How a home robot interacts with people is one of most important issue to be accepted by people who want to share their time and spaces with robots.

Various researches have been studied in a field of human-robot interaction. These studies are able to classify into two groups of interaction. The first is the group of using direct commanding method between a human and a robot. Most of the researches deal with methods of communication such as gesture [3], [4], [5], speech [6], [7], [8] and using control devices like joysticks or computers [9], [10], [11]. We call this type of interaction DCM (Direct Commanding Method). Fig. 1 shows DCM interaction model. In this model, a human has two tasks: to control a robot by commanding and to act to environment. For example, in the scenario of using a cleaning robot, it can sweep the region pointed out by a human. The human has to point out a region and to remove obstacles in order for the robot to move about in the environment. In this case, interactions between a human and a robot are described as follow (a human, a robot and environment are expressed as H, R and E respectively) :

- $H \Rightarrow R$: the human points a cleaning region by gesturing.
- $R \Rightarrow H$: the robot indicates acknowledgement of a gesture command.
- $H \Rightarrow E$: the human removes obstacles.
- $E \Rightarrow H$: state of the environment is perceived.
- $R \Rightarrow E$: the robot cleans a region and moving about in the environment.
- $E \Rightarrow R$: state of the environment is sensed.

In contrast, the second group has no explicit commanding. There are studies related human-robot cooperation such as the cooperation of carrying a long or big object by a human and a robot based on a manipulator [12], [13], [14], and outdoor cooperative tasks by a human and a humanoid [15]. We call this type of interaction CEA (Commands Embedded in Actions). Fig.2 shows CEA interaction model. CEA has no direct interaction between a human and a robot because the human can control the robot by executing his/her actions to environment. For example, in the case of carrying a long object by a human and a manipulator robot, its interaction is described as follow:

- $H \Rightarrow E$: the human moves about in the environment.
- $E \Rightarrow H$: state of the environment is perceived.
- $R \Rightarrow E$: the robot moves about in the environment.
- $E \Rightarrow R$: state of the environment is sensed.

The robots can work for helping a human by sensing force of the shared object without DCM. The human does not need to execute direct commanding to a robot and understand a way for communication with it.

CEA reduces a human work-load because it needs no additional action and no understanding a way for communication. In fact, we can find its applications of CEA such as automatic doors, automatic faucets and so on easily, and they are certainly convenient. However, these applications deal with easier commands because tasks of automatic doors are not so complicated. We then consider CEA needs to be extended for more complicated tasks, and propose ECEA (Extended CEA). In this paper, we investigate an application of ECEA about cooperative sweeping by a human and a mobile robot.

In the rest of this paper, we describe the detail of Extended

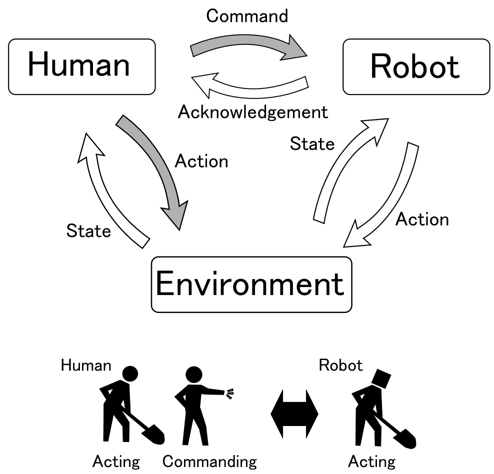

Fig. 1. DCM Model

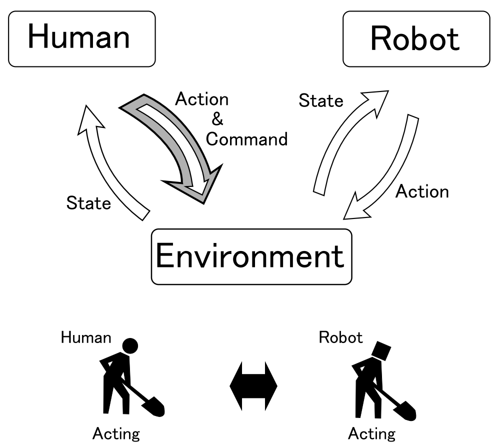

Fig. 2. CEA Model

CEA and its specific application, cooperative sweeping in Section 2 and 3 respectively. In Section 4, we describe experiments to compare ECEA with DCM in terms of human's cognitive loads. We discuss the experimental results in Section 5, and conclude in Section 6.

II. Extension of Commands Embedded in Actions

Before describing the extension of CEA, we explain Action Coding System (ACS) [16]. The ACS is the key method of embedding commands into human actions.

A. Action Coding System

We employ ACS to decompose a human action into primitive acts and to embed commands into his/her action. The ACS provides a general description of the events realized by a human in carrying out the actions of simple activities of daily living tasks. It was used for analyzing the brain damaged patients' behavior [16]. The ACS describes two levels of action units which are defined as follow:

- *A-1s:* A-1 units are the smallest components of a behavioral sequence that achieves a concrete, functional result or transformation, describable as the movement of an object form one place to another or as a change in the state of an object.
- *A-2s:* A-2 unit is a group of part of A-1s that accomplish one of basic subgoal of the task.

(1) MOVE (x) TO (location) VIA (instrument) BY (manner)	
(2) ALTER (x) TO (location) VIA (instrument) BY (manner)	
(3) TAKE (x)	(i.e. take possession of object x)
(4) GIVE (x)	(i.e. relinquish possession of x)

Fig. 3. Four basic A-1 types

- **Opening sugar pack:**
  ALTER sugar pack TO open BY tearing
- **Pouring sugar:**
  MOVE sugar TO in coffee VIA pack BY pouring

Fig. 4. Example of coffee making task

For example, Schwartz et al. [16] coded four A-1 types showed in Fig.3 by reviewing 28 breakfast videos. All MOVES and many ALTERS contain within them the taking and relinquishing of possession of objects. To minimize complexity and avoid multiplying action descriptions, they coded TAKE(x) and GIVE(x) only when the change of possession was temporally discontinuous with the action carried out on x. Using this notation, specific acts are expressed in Fig.4. In our study, we employ the A-1s as primitive units of human action.

B. Extension of CEA

Using the ACS, we can clarify the structure of CEA. In the case of using an automatic door, the action of "ALTER door To open" is embedded in a human A-1 unit of "MOVE one's body To beyond the door" as a command. The "ALTER" is performed by a machine instead of the human. However, this is one-shot interaction: only one robot's action similar to an A-1 unit corresponds to a human A-1 unit. It is difficult to embed multiple commands into a human A-1 unit by CEA. We then consider CEA has the limitation for performing more complicated tasks.

The idea of ECEA is that the changes of human actions with keeping his/her work-load low and increases A-1s where commands can be embedded. The definition of ECEA's guideline is described as follow:

- To minimize spatial change of a human action.
- To minimize temporal change of a human action.

They minimize the differences in trajectories or in time between an original human action and an extended one. This kind of extension is realized by embedding commands between A-1 units. It means a relaxation of the strength of a connection between A-1 units. However ACS provides candidates of the point where command are embedded, it cannot choose the optimal one because such a point significantly depends on tasks.

- Original action sequence:
TAKE sugar pack
ALTER sugar pack TO open BY tearing

- After applying temporal extension:
TAKE sugar pack
KEEP taking (keep final state of taking)
ALTER sugar pack TO open BY tearing
KEEP altering (keep final state of altering)

Fig. 5. Example of temporal extension

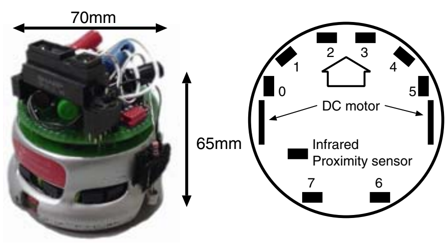

Fig. 6. KheperaII

The new command "KEEP" is employed. It keeps a final state of a previous act. An example of the "KEEP" is described in Fig.5. Such extension follows our guideline. To confirm effectiveness of ECEA, we apply ECEA for a specific task, a cooperative sweeping between a human and a robot, and perform experiments to measure human cognitive loads in the task.

III. Design on Cooperative Sweeping

We deal with a human-robot cooperative sweeping by a human and a small mobile robot as a cooperative work. A goal of the cooperative task is to sweep out a desk including the region of under an object. In this section, we first describe an experimental environment and a specification of small robot. Next, we describe A-1 scripts of this task and apply ECEA.

A. Environment and robot

We assume the environment where a human and a robot work cooperatively is a place used by a human routinely such as a desktop. A desk swept out by a robot has a flat surface and a wall which encloses the region for keeping a robot not to fall. We use a small mobile robot KheperaII (Fig.6). The robot has eight infrared proximity and ambient light sensors with up to 100mm range, a processor Motorola 68331 (25MHz), 512 Kbytes RAM, 512 Kbytes Flash ROM, and two DC brushed servo motors with incremental encoders. The program written by C-language runs on the RAM.

B. Tasks

We assign tasks for a human and robot with considering competence of the mobile robot, a work-load for human. Assigned tasks for a human and a robot are described respectively:

- A robot's tasks are to sweep out a desk autonomously with strategy of a random turn and to sweeps out the region of under an object when a human moves the object.
- A human's task is to move an object in the environment.

Our robot can only use local information which is obtained by its sensors and cannot move objects by its hardware equipment. It is difficult to employ effective region covering[17] with creating map because it needs a correct position of a robot. Our robot cannot obtain its correct position by its sensors (e.g. a dead reckoning method has low reliability because of its accumulated errors).

To reduce a human work-load, it is effective to sweep out the region under an object when a human moves the object. However, it is difficult for our robot to sweep there automatically because it has no competence to detect the region of under an object. A human needs to control the robot by multiple explicit commands using DCM. CEA also cannot apply for this because it cannot deal with multiple commands. We therefore extend CEA and create A-1 scripts of human actions. The human has two actions to achieve his/her task: (1) *pick up an object and put down it after cleaning* and (2) *move an object to another place and move it to an initial location after cleaning*. In the action of (1), we can see that the robot can sense the edge of the object above it by its local sensors and sweep out the region of under the object while a human picks up it. The robot repeats to turn when it senses the edge of an object, and then the sweeping is performed.

We then introduce a A-1 unit "KEEP" which keeps a final state of a previous act. The A-1 units scripts of an original human action and an extended one are described in Fig.7 and Fig.8 respectively. The "KEEP" in the extended action is originally included in the "MOVE", and it is not completely new action for a human. In this point, we consider a human has no additional load. As a result of the extension, multiple commands are able to embedded while the "KEEP" is continued. This is a specific instance of ECEA and we call this temporal extension.

IV. Experiments

We conduct experiments to confirm that our interaction design reduces work-load of human. We examine human cognitive load to evaluate human work-load, and compare the load between ECEA and DCM. The experiments are performed on our experimental system described in the next subsection.

A. Experimental system

Fig.9 shows the experimental system which can indicate a robot's trajectory. This system consists of a sweeping area and a projection system. In experiments, a human interacts with a robot on the sweeping area in Fig.10 indicating a swept location by the projection system including a personal computer, a projector, and a USB camera.

The projection system detects a robot's location using a picture of two beams of infrared LEDs equipping on the robot.

(1) TAKE object
(2) MOVE object TO z (z: a vertical location)
(3) MOVE object TO x (x: an initial object's location)
(4) GIVE object

Fig. 7. Typical acts of moving the object

(1) TAKE object
(2) MOVE object TO z
(2') KEEP moving (keep final state of moving)
(3) MOVE object TO x
(4) GIVE object

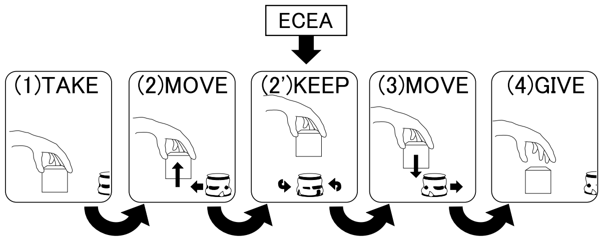

Fig. 8. Applying ECEA with temporal extension

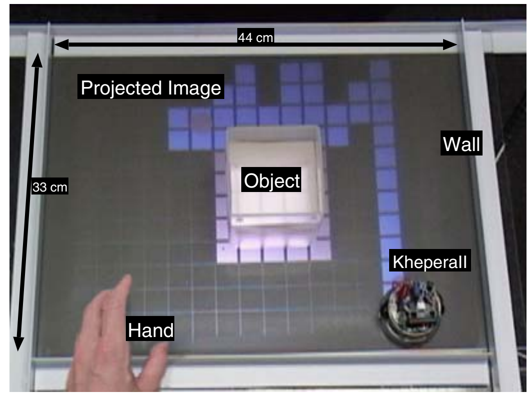

Fig. 10. Sweeping area

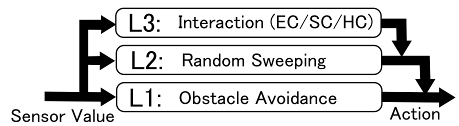

Fig. 11. Subsumption architecture

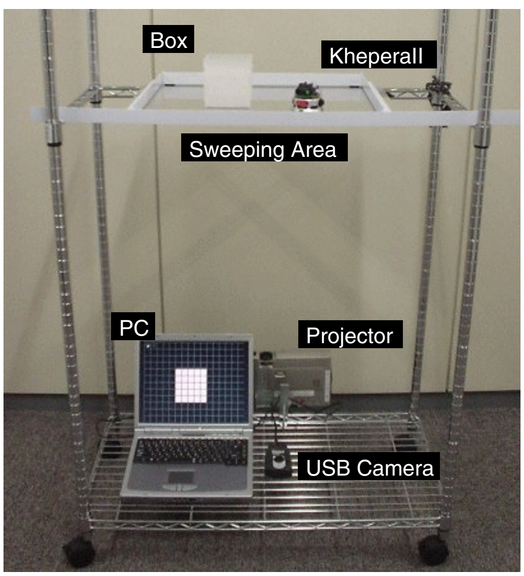

Fig. 9. Experimental system

The robot's location is calculated by image processing in the picture from the camera, and then an image indicating the trajectory is created with the location. This image is ultimately projected on the sweeping area.

The projected image also includes small square cells. A cell is lit when a robot enters the cell in real time. These small cells therefore express the trajectory of a robot. The sweeping area having a width of 44 cm and a height of 33 cm is divided into 16×12 cells. Cells of 3×3 approximately correspond to the area of a robot.

B. Behavior Design of Mobile Robot

A robot is implemented by behavior-based approach, and we adopt subsumption architecture [18] to manage following behaviors:

- *L1 Obstacle Avoidance:* a robot stops when it approaches obstacles, and it goes backward when an obstacle approaches it.
- *L2 Random Sweeping:* a robot going forward when no obstacles are on its front, and turning for random direction when it senses obstacles such as wall.
- *L3 Interaction:* a robot turns randomly when an object is above it or when it receives a command from a human.

Fig.11 shows the robot's behaviors into the three layers in subsumption architecture. Each layer asynchronously checks the applicability of behaviors and executes applicable ones. Higher layers suppress lower layer's behaviors, and lower layers have more reactive behaviors. The behaviors of each layer consist of multiple actions. When the system obtains multiple outputs, it generally selects the highest layer's action. Each layer has output frequency of action to control the robot smoothly. We set the frequency as obstacle avoidance: an action by 5msec, sweeping: 10msec, interaction: 5msec, obstacle avoidance and interaction occur most frequently.

C. Cognitive load measurement

We measure the cognitive load of a human interacting with a robot by ECEA and DCM for comparison. Two methods are chosen as typical DCM without extra devices such as remote controllers. These control methods are shown in Fig.12, and the detail is described as follow:

EC : A robot performs sweeping by ECEA.

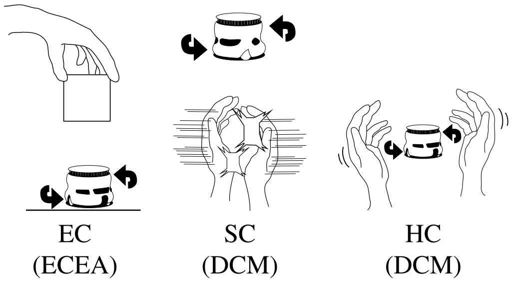

Fig. 12. Three types of interactions

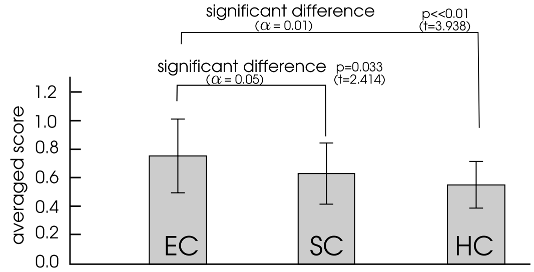

Fig. 13. Results of scores and differences

SC : A robot sweeps with receiving a sound command by hand clapping.

HC : A robot sweeps with receiving a command by hand as blocking the robot's line.

The robots receiving such commands are prepared by adding extra sensors such as microphone or modifying the program of a robot. When a robot controlled by DCM receives a command from a human, it turns from $90°$ to $180°$ randomly. A robot controlled by ECEA also turns from $90°$ to $180°$ randomly when it senses the edge of the object over its head.

Measurement starts when a robot enters the region of under an object, and it continues until all cells of the region are swept. A box whose size is 4×4 cells is employed because sweeping time of the region is appropriate for subjects and the measurement. In the EC, subjects keep to pick up an object until all the cells of under the box are swept. In the SC and HC, first, a human relocates a box to a corner of the sweeping area, and then send a command for a robot to be turn by making sounds or approaching their hand to it when it is likely to run out from the region of a box.

We use a dual task method to measure human cognitive load. Subjects have to do mental arithmetic as a secondary task while controlling the robot as a primary task [9], [19]. They count backwards by three from a randomly selected three-digit number vocally. We obtain the average number of correct answers per second, and evaluate the human's cognitive load for controlling each robot. Subjects are required to calculate as quickly as possible, and to prioritize the controlling a robot over the counting. They practice controlling robots and the counting well before experiments begin. The order of EC, SC and HC for each subject is determined at random, and these experiments are recorded three times respectively for a subject. Subjects are also measured counting ability without operations of a robot before a measuring of EC, SC, or HC respectively. The counting ability is the number of correct answers of the counting for 30 seconds.

D. Results

Subjects are eight men and four women between the age of 22 and 32. Each subject has three scores: EC, SC, and HC. A score is the average of normalized number of correct counting answers per second for a subject. The normalization is to divide the correct answers per second by correct answers per second without operations of a robot. Fig.13 shows averaged scores, standard deviations, and differences tested by Dunnett's test. Each EC, SC and HC is the average of all subject's scores. The number 1.0 means counting ability of each subject without operations of a robot. EC has the highest average. The difference between EC and HC has a significant difference ($p << 0.01, \alpha = 0.01, t = 3.938$). The difference between EC and SC also has a significant difference ($p = 0.033, \alpha = 0.05, t = 2.414$). Fig.14 shows the experimental appearance.

The results of the experiments show that the ECEA reduces a human cognitive load in comparison with other DCM. ECEA has a low cognitive load because of minimizing cost of sending commands and also shortening the trajectory of moving a box. Therefore, ECEA plays the significant role in a human-robot cooperative task.

V. Discussion

Physical loads appear to influence the cognitive loads. Each DCM accompanies motions of human arms. Hence, the measured cognitive loads would include a load of the motions and a load of human attention. However, we consider that the effects of human motions on the measured cognitive loads keep a minimum because the most natural and intuitive form of task performance by a human is to perform the task in the usual manner: using his/her own hands. Actually, the subjects have no choice except to clap his/her hands or to move his/her arms in the experiments. Therefore, the difference between ECEA and DCM is attention for a robot. A subject has to repeat the cycle of observation of a robot and execution of moving his/her arms in the experiments of SC and HC. In contrast, in the experiments of EC, a subject does not need to concentrate his/her attention on a robot, and he/she can interact with environment naturally like normal CEA in Fig.2. In addition, the experiments are set to be fair deal between ECEA and DCM in terms of controlling a robot without specific devices and the robot's software codes for its behavior are almost same among EC, SC and HC.

An advantage of the CEA/ECEA model over conventional models of plan-recognition-based cooperation is that it does not need complicated and computational costly plan recognition process. For example, in Grosz and Sidner's SharedPlan[20], a robot needs to infer user's beliefs and

Fig. 14. Experimental Appearance

intentions. However, it has a disadvantage that it is not able to change the cooperation plan when the situation is changed.

VI. CONCLUSION AND FUTURE WORK

We first classified present human-robot cooperation into two groups of interaction: DCM (Direct Commanding Method) such as gesture commanding method and CEA (Commands Embedded in Actions) which a human could control a robot by through execution of his/her actions to environment. We proposed ECEA (Extended CEA) in order to deal with more complicated tasks than CEA. On the cooperative sweeping task between a human and a mobile robot, we applied temporal extension as one of ECEA instances to the human action. Multiple commands were embedded in the action by the extention and the robot performed more complicated task with the human. The experiments for conforming reduction of a human work-load using ECEA were conduced on the sweeping task. Human cognitive loads were measured as human work-loads and compared between ECEA and DCM. The results of the experiments showed that the ECEA minimized a human cognitive load. We therefore confirmed that our temporal extension was able to apply more complicated task than CEA and caused lower cognitive load than DCM.

We currently investigate other instances of ECEA and apply them to many practical tasks such as daily living tasks and more complicated tasks with considering A-2 units of action coding system. We are planning to examine another aspect of ECEA's advantages. For example, a human might notice the functions of a robot easier by using ECEA. In addition, to use higher level embodied robots like AIBO is also necessary.

REFERENCES

[1] G. S. Hornby, S. Takamura, J. Yokono, O. Hanagata, T. Yamamoto, and M. Fujita, "Evolving robust gaits with aibo," in *Proc. of IEEE International Conference on Robotics and Automation (ICRA'00)*, 2000, pp. 3040–3045.

[2] W. Burgard, A. B. Cremers, D. Fox, D. Hahnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun, "The interactive museum tour-guide robot," in *Proc. of the Fifteenth National Conference on Artificial Intelligence (AAAI/IAAI)*, 1998, pp. 11–18.

[3] J. Triesch and C. von der Malsburg, "Robotic gesture recognition," in *Proc. of the Bielefeld Gesture Workshop*, 1997, pp. 233–244.

[4] S. Waldherr, R. Romero, and S. Thrun, "A gesture based interface for human-robot interaction," *Autonomous Robots*, vol. 9, no. 2, pp. 151–173, 2000.

[5] F. Marrone and M. Strobel, "Cleaningassistant - a service robot designed for cleaning tasks," in *Proc. of Advanced Mechatronic Systems (AIM'01)*, 2001.

[6] C. Breazeal and L. Aryananda, "Recognition of affective communicative intent in robot-directed speech," *Autonomous Robots*, vol. 12, no. 1, pp. 83–104, 1 2002.

[7] S. Lauria, G. Bugmann, T. Kyriacou, and E. Klein, "Mobile robot programming using natural language," *Robotics and Autonomous Systems*, vol. 38, no. 3–4, pp. 171–181, 3 2002.

[8] S. Kajikawa, S. Hiratsuka, T. Ishihara, and H. Inooka, "Robot position control via voice instruction including ambiguous expressions of degree," in *Proc. of IEEE Int. Workshop on Robot and Human Interactive Communication (ROMAN'03)*, 2003.

[9] J. W. Crandall and M. A. Goodrich, "Characterizing efficiency of human-robot interaction: A case study of shared-control teleoperation," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'02)*, 2002.

[10] D. Katagami and S. Yamada, "Active teaching for an interactive learning robot," in *Proc. IEEE Workshop Robot and Human Interactive Communication (ROMAN'03)*, 2003.

[11] A. M. Khamis, F. J. Rodríguez, and M. A. Salichs, "Remote interaction with mobile robots," *Autonomous Robots*, vol. 15, no. 3, 2003.

[12] Y. Hayashibara, Y. Sonoda, T. Takubo, H. Arai, and K. Tanie, "Assist system for carrying a long object with a human – analysis of a human cooperative behavior in the vertical direction –," in *Proc. of 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'99)*, 1999.

[13] H. Arai, T. Takubo, Y. Hayashibara, and K. Tanie, "Human-robot cooperative manipulation using a virtual nonholonomic constraint," in *Proc. of 2000 IEEE International Conference on Robotics and Automation (ICRA'00)*, 2000.

[14] K. Nakai, K. Kosuge, and Y. Hirata, "Control of robot in singular configurations for human-robot coordination," in *Proc. of IEEE Int. Workshop on Robot and Human Interactive Communication (ROMAN'02)*, 2002, pp. 356–361.

[15] K. Yokoyama, J. Maeda, T. Isozumi, and K. Kaneko, "Application of humanoid robots for cooperative tasks in the outdoors,," in *Proc. of IEEE/RSJ IROS Workshop on Explorations towards Humanoid Robot Applications*, 2001.

[16] M. F. Schwartz, E. S. Edwards, M. Montomery, C. Palmer, and N. H. Mayer, "The quantitative description of action disorganization after brain damage: A case study," *Cognitive Neuropsychology*, vol. 8, pp. 381–414, 1991.

[17] H. Choset, "Coverage for robotics - a survey of recent results," *Annals of Mathematics and Artificial Intelligence*, vol. 31, pp. 113–126, 2001.

[18] R. A. Brooks, "A robust layered control system for a mobile robot," *IEEE Journal of Robotics and Automation*, vol. 2, no. 1, pp. 14–23, 1986.

[19] J. E. Condron and K. D. Hill, "Reliability and validity of a dual-task force platform assessment of balance performance : Effect of age, balance impairment, and cognitive task," *Journal of American Geriatrics Society*, vol. 50, pp. 157–162, 2002.

[20] B. Grosz and C. Sidner, "Plans for discourse," in *Intentions in Communication*, P. Cohen, J. Morgan, and M. Pollack, Eds. MIT Press, 1990, pp. 417–444.