2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

Task instruction by putting task information in work space

Kazuyuki NAGATA, Yujin WAKITA, and Eiichi ONO
National Institute of Advanced Industrial Science and Technology (AIST)
Intelligent Systems Research Institute
1-1-1 Umezono, Tsukuba, 305-8568, Japan
Email: k-nagata@aist.go.jp

Abstract— **This paper presents a task instruction scheme for a service robot operating in a daily life environment. It is difficult for the robot to execute tasks automatically in the daily life environment, because it should handle various objects and the working environment is complex. In such an environment, the concept of shared autonomy by which the robot can share a task with a human is important. The task instruction that we propose is as follows: An operator clicks on a target object in a camera image and the system gets the position of the clicked point in the work space from the corresponding range image. Then, the operator selects the task model and attaches the task information to the position and the robot can perform its task safely and reliably using this task information. The features of the task instruction are as follows. An object model is composed of a list of task models which can be applied to the object category. The user recognizes the object type and its situation in the environment, and instructs a robot to execute a task by selecting a task model according to the object type and situation.**

I. INTRODUCTION

In recent years, service robots have been studied in many areas of application, such as home and office as supports for human daily life. The service robot must execute its task with absolute safety and should handle many kinds of objects in a daily life environment. Therefore, it is difficult for a robot to execute all tasks automatically in such an environment. Of course, a robot can execute all tasks in a daily life environment by tele-operation, but this leaves a large part of the burden of the operation to the user.

An alternative has been proposed with the concept of shared autonomy [1]. In shared autonomy, the robot and the human operator cooperate to achieve a task by sharing the parts of this task. Each part of the task is suitable for the robot and the human. During the robot's autonomous task execution, the robot can execute its own task and can require the human operator's help for the part of the task that is hard for it to execute. The concept of shared autonomy is important in the task instruction to the robot.

Among the tasks assigned to service robots, "Picking up an object" is a simple task but a very important one, and a lot of research has been done on picking up an object in a daily life environment [2]-[8]. Miura et al. studied a service robot in which the robot visually recognizes the target object and picks it up [2], and Yamazaki et al. developed a mobile robot which can acquire an object model through a video image and manipulate this object [4]. However, it is difficult for the robot to visually recognize all objects autonomously in such a daily life environment because appropriate lighting condition is required for visually recognizing objects and objects can be occulted by other ones, which is a severe problem in the daily life environment. ID tags are used to identify the object that the service robot has to manipulate [5][6], but ID tags information is not really appropriate for manipulating the object. Task models are necessary to manipulate objects, which may be changed according to the situation, for example, whether the object is standing, fallen, hidden by another one and so on and thus the object information written in ID tags may not be sufficient to manipulate it. In another research on task instruction for a service robot, Saito et al. developed a teaching tablet interface, in which the user draws a figure in the manner of 3D CAD on the camera image of the remote environment to teach task information to the robot [7].

We think that task instruction should be simple for a service robot to execute simple task such as picking up an object in daily life environment. And we think that the complexity of daily life environment should be considered to instruct a robot to execute tasks and this complexity can be defined as follows:

- Multiple Objects:
  There are many kinds of objects in a daily life environment and even for a single kind of object, these objects can have different sizes and designs. Therefore, it is difficult for the robot to have individual object models for everyday object.
- Temporal Complexity:
  In daily life environment, not only the robot but also the human may move the object and the robot may not recognize the change of the object situation. So the robot cannot know the object position and situation perfectly before executing its task.
- Spatial Complexity:
  The object may be put in a narrow space such as in the drawer of a desk or on a shelf and other objects may surround the target object. Such situations can affect the robot motion.

In this paper, we present a novel task instruction system

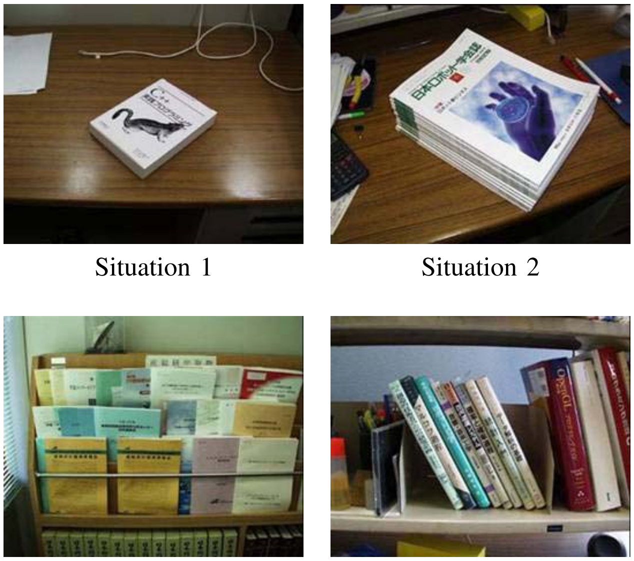

Fig. 1. Various types of book arrangements

for a service robot considering the complexity of the daily life environment. Firstly the models for the task instruction are described, then the basic setup of the system and the procedure of the task instruction are described and finally experimental results are presented.

II. MODELS FOR TASK INSTRUCTION

The strategy for picking up an object with a robot is different according to the object type and situation in the environment. For example, Fig. 1 shows various types of books placed in various positions. In situation 1 only one book is placed on a desk, in situation 2 books are piled up on a desk, in situation 3 books are placed on a magazine rack, and in situation 4 the books are placed on a book stand. The task strategy and task information needed for picking up a book is different according to the type of book arrangement. So one needs to know the situation of the object placed in the environment for picking it up. Furthermore it is difficult for a robot to identify an object and to automatically recognize how it is placed in the environment under unknown lighting condition and when two or more objects are placed in a disorderly manner such as in the daily life environment. In contrast, a human can instantaneously identify an object and recognize the object situation in the environment. In our task instruction system, the user identifies the object and recognizes the situation of the object placed in the environment, and instructs a robot to execute the task by selecting the task model according to the object type and situation.

A. Task Model [9]

A task model describes the knowledge of the task, i. e. task parameters which are necessary to execute the task and a template of the sub-task sequence. The sub-task describes

Fig. 2. Various cups

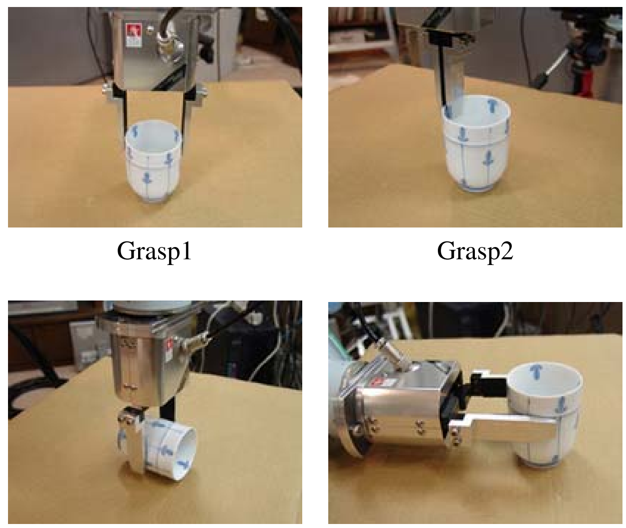

Fig. 3. Four types of cup grasping with a robot gripper

the acquisition of the task parameters and a primitive motion that comprises the task. During the sub-task execution, the robot can execute its own sub-task and can require the human operator's help for the part of the sub-task that is hard for it to execute. That is, the robot can ask more information to the human operator for getting more task parameters [10][11], and the robot can ask the operator to control it manually when it cannot execute the automatic motion.

B. Object Template Model (OTM)

We are surrounded by various objects, however, handling objects belonging to the same category has common features even though they have different sizes and designs. For example, Fig. 2 shows various cups and tasks of grasping a cup with a robot gripper are classified into four types as shown in Fig. 3. If the robot has task models for picking up a cup of these 4 grasping types, it can pick up various cups in various situations. Various objects that exist around us are classified from the viewpoint of object handling. We define the Object Template Model (OTM) that is the list of task models which can be applied to an object category.

The objects which belong to the same object category are expressed by one OTM. Therefore, the robot does not need to have an individual object model. The operator instructs the robot to execute a task by selecting the task model described in the OTM according to the object situation.

C. Task Space Model (TSM)

Information on an object fixed in environment such as floor, wall, and furniture can be described in the system beforehand. We define the Task Space Model (TSM) which describes spatial information of an object fixed in an environment related to tasks of the robot. The TSM contains geometrical information such as the height of a desk, information that relates to the robot motion such as approach directions to the target object, and task skills such as opening a door or a drawer, etc. Moreover, when the user customarily puts specific objects on a particular place such as on a kitchen cabinet, or a bookshelf, the OTMs of the objects can be included in the TSM. Our work space model is constructed by arranging the TSMs in the system according to the arrangement of the real object fixed in the environment.

III. TASK INSTRUCTION

A. Basic Setup of the System

The basic setup of the task instruction system is a manipulator and a stereo camera. The camera image of the work space is displayed on a screen and a range image of the work space is captured by the stereo camera. The stereo camera has it's own coordinate frame. The work space model is described in the coordinate frame of the stereo camera. Therefore the TSMs in the work space are also described in the coordinate frame of the stereo camera. The system has the OTMs for object categories which the robot may handle, and task models which may be executed by the robot are described in the OTMs. Furthermore, the OTMs are included in the TSM according to the user's habits.

B. Procedure of Task Instruction

The procedure of the task instruction is as follows,

1) A user finds a target object in a camera image and clicks on it (Fig. 4). The user points at the target object directly in this way, so that the robot does not need to search for it.

2) The system gets the position of the clicked point on the target object in the work space from the corresponding range image and selects the TSM in which the clicked point is included. The information of the TSM is attached at the position of the clicked point.

Fig. 4. The user clicks on a target object in a camera image

Fig. 5. Example of icon list of the Object Template Model (OTM)

Fig. 6. Example of icon list of the task model

3) The icon list of the OTMs which are included in the TSM is displayed (Fig. 5).

4) The user identifies the target object and selects the proper OTM from the icon list, and clicks on the icon. If there is no proper OTM in the icon list, the user accesses the OTM-library and selects the proper OTM.

5) The system displays the icon list of the task models which are described in the selected OTM (Fig. 6).

6) The user selects the task model while considering the object situation in the work space, and clicks on the icon of the task model. If there is no proper task model in the icon list, the user accesses the task model library and selects the proper task model. The selected task model is attached at the position of the clicked point in the work space.

Fig. 7. The system asks a user to input additional information

7) The robot executes the task at the position of the clicked point in the work space using the space information which is described in the TSM and the task model. If there is no proper task model, the user controls the robot manually.

IV. EXPERIMENT

We conducted experiments on a task instruction in which we instruct a robot to pick a receptacle for chemical experiment which was placed on a desk. The picking task is comprised of the following 7 primitive motions, 1. move-to-approach, 2. pre-grasp, 3. approach, 4. grasp, 5. lift-up, 6. departure, 7. move-to-destination. A manual operation module for robot control was inserted between each primitive motion to correct the robot motion.

Details of these primitive motions are as follows,

- move-to-approach:
  This motion moves the robot hand to the starting point of the approach motion. The parameters which are necessary for this motion are position and orientation of the robot hand at the approach starting point. These parameters are calculated by using the information on the approach direction, approach distance and position and orientation of the robot hand at the grasping point. In our experiments, the approach direction was given from the TSM and the approach distance was set at a fixed value.

- pre-grasp:
  The robot hand opens for grasping a target object.

- approach:
  This motion is that of the robot hand moving to the grasping point in low speed. The position and orientation of the hand at the grasping point is needed for this motion.

- grasp:
  The robot hand grasps a target object.

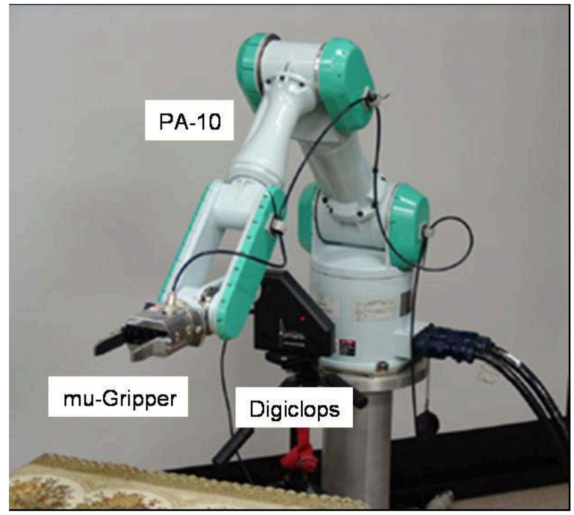

Fig. 8. Experimental setup

- lift-up:
  This motion is that of the robot lifting up a grasped object. This motion needs information on the moving distance. In the experiments, the moving distance was set at a fixed value.

- departure:
  This motion is that of the robot hand moving to the safety area. The parameter giving the departure point is necessary for this motion. In the experiments, the departure point was set at a fixed value.

- move-to-destination:
  This motion is that of the robot hand moving to a destination point. The parameters giving the position and orientation of the robot hand at the destination point are necessary for this motion. In the experiments, the position and orientation at the destination point were set at a fixed value.

If the system gets the parameters of the position and the orientation of the robot hand at the grasping point, then all information necessary for this task can be calculated. The task of acquiring the parameters for position and orientation at the grasping point was carried out at first. The position and orientation at the grasping point were calculated by using the information of the position which was acquired by the initial click on the target object and the initial position of the manipulator. If the system needs more information such as in the case of an object lying on a desk, it asks the user to input the additional information (Fig. 7). After getting the information, the robot executes the primitive motion sequentially from move-to-approach to move-to-destination.

As shown in Fig. 8, the experimental setup is composed of a PA-10 (Mitsubishi Heavy Industries, LTD.) robot arm, mu-Gripper (Takano Bearing Co. , LTD.) robot hand, and a Digiclops (Point Grey Research Inc.) stereo camera placed

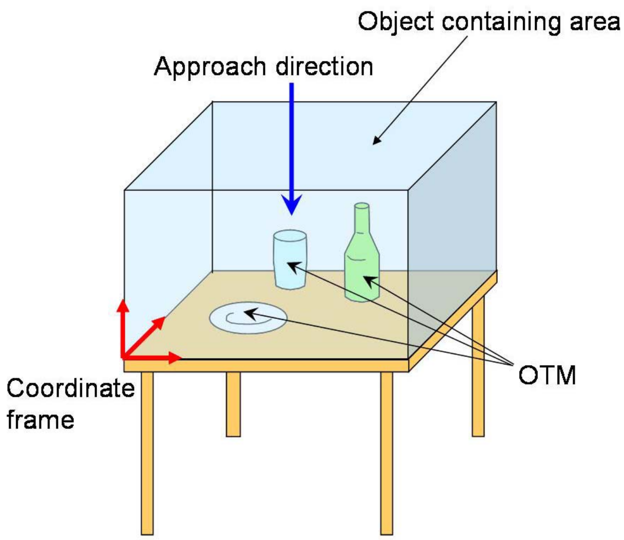

Fig. 9. Task space model of a desk

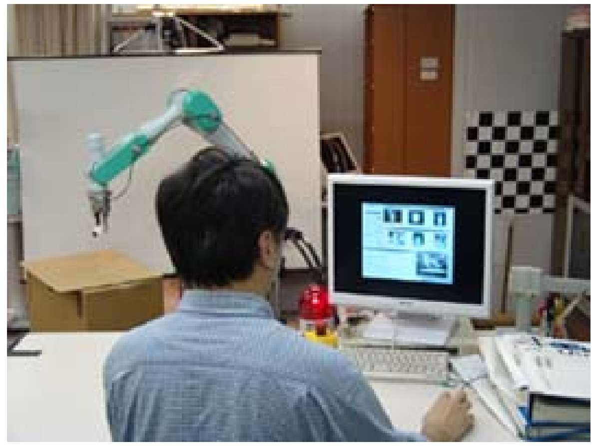

Fig. 10. Task instruction scene

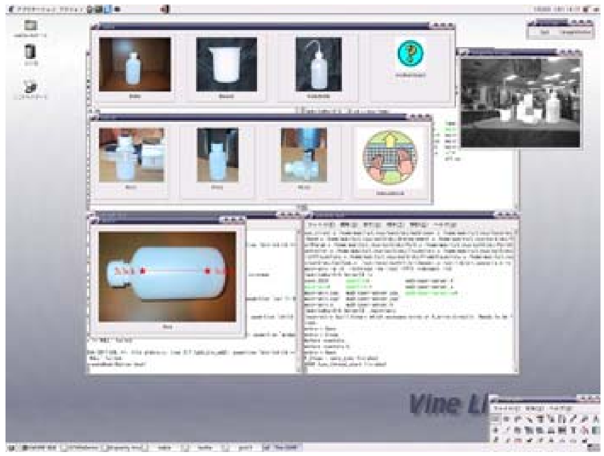

Fig. 11. Screen image

beside the manipulator. The coordinate frame of the stereo camera with respect to the coordinate frame of the manipulator was calculated by calibration. The Digiclops camera gives a range image of the work space. As shown in Fig. 9, the coordinate frame of the desk with respect to the coordinate frame of the stereo camera, the area containing target objects which we call the object-containing-area, and the approach direction to the target object are described in the TSM of the desk. The approach direction to the target object was set to the vertical direction from top to bottom. In the experiments, we assumed that a receptacle for a chemical experiment is placed on a desk, so the OTMs of a bottle, a beaker, and a water bottle were included in the TSM of the desk.

First of all, an operator clicked on a target object in a camera image (Fig. 4). The 3D position of the clicked point in the work space was given to the system by the Digiclops. Next, the system checked whether the clicked point is inside the object-containing-area of the TSM of the desk. When the clicked point was inside the object-containing-area, the information which is described in the TSM was attached at the point. Then, the icon list of the OTMs included in the TSM was displayed (Fig. 5) and the operator selected the bottle of the OTM and clicked on the bottle icon. Then the icon list of task models which are included in the OTM of the bottle was displayed (Fig. 6) and the operator selected the picking task on the right of Fig. 6 and clicked on the icon. In this case, an information on direction of the target object is needed to calculate the orientation of the hand at the grasping point. The system displayed a window for input additional information (Fig. 7) and asked the operator to click two points on the object in the camera image. Then, the robot executed the picking task using the information on the clicked point and selected task model.

Fig. 10 shows task instruction scene, Fig. 11 shows the screen image during task instruction, and Fig. 12 shows the experimental scenes. Fig. 12 (a) shows initial condition, (b) shows the robot starting approach motion, (c) shows the robot reaching the grasping point, (d) shows the robot grasping the target object, (e) shows the robot lifting up the object, and (f) shows the robot reaching the departure point.

V. CONCLUSION

This paper presents a novel task instruction system for a service robot operating in the daily life environment. Firstly, we classified various objects that exist around us in the daily life environment from the viewpoint of object handling, and we defined the Object Template Model (OTM) that is the list of task models which can be applied to the object category. We also defined the Task Space Model (TSM) which describes a list of OTMs and spatial information of an object fixed in an environment related to tasks of the robot. The user can put the OTMs in the TSM according to his/her custom. Our work space model is constructed by arranging the TSMs in the system according to the arrangement of the real object fixed in the environment. The proposed

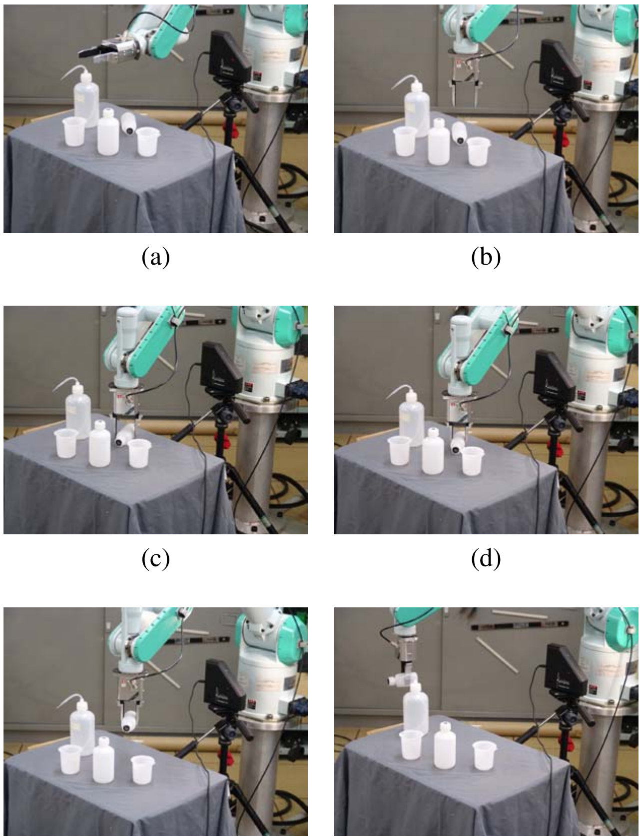

Fig. 12. Experimental scenes

task instruction system is composed of a manipulator and a stereo camera. The camera image of the work space is displayed on a screen and a range image of the work space is captured by the stereo camera. The stereo camera has it's own coordinate frame. And the work space model is described in the coordinate frame of the stereo camera. A user finds a target object in a camera image and clicks on it with a mouse button. The system gets the position of the clicked point in the work space from the corresponding range image and the space information is attached at the position from the TSM. The user identifies the target object and selects the OTM described in the TSM. The user then selects the task model while considering the object situation in the work space. Then the robot executes the task at the position of the clicked point in the works space using the spatial information which is described in the TSM and the task model attached at the position.

The two features of the task instruction are as follows. First, an object model is composed of a list of task models which can be applied to the object category. Second, the user recognizes the object type and its situation in the environment in place of the robot because the robot is not good enough at the recognition of the real world. The user's recognition is transferred to the robot by the simple selection of the OTM and of the task model.

References

[1] S. Hirai, T. Sato and T. Matsui, "Intelligent and Cooperative Control of Telerobot Tasks," *IEEE Control Systems*, Vol. 12, No. 1, pp. 51–56, 1992.

[2] Y. Makihara, M. Takizawa, K. Ninokata, Y. Shirai, J. Miura and N. Shimada, "A Service Robot Acting by occasional Dialog - Object Recognition Using Dialog with User and Sensor-Based Manipulation -," *Journal of Robotics and Mechatronics*, Vol. 14, No. 2, pp. 124–132, 2002.

[3] J. Suthakorn, S. Lee, R. Thomas, S. Choudhury and G. S. Chirikjian, "A Robotic Library System for an Off-Site Shelving Facility," *Proc. of the 2002 IEEE Int.@Conf. on Robotics and Automation*, pp. 3589–3594, 2002.

[4] K. Yamazaki, M. Tomono, T. Tsubouchi and S. Yuta, "A Grasp Planning for Picking up an Unknown Object for a Mobile Manipulator," *Proc. of the 2006 IEEE Int.@Conf. on Robotics and Automation*, pp. 2143–2149, 2006.

[5] N. Y. Chong and K. Tanie, "Object Directive Manipulation through RFID," *Proc. of on Control, Automation, and Systems*, pp. 22–25, 2003.

[6] N. Y. Chong, H. Hongu, K. Ohba, S. Hirai and K. Tanie, "A Distributed Knowledge Network for Real World Robot Applications," *Proc. of 2004 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 187–192, 2004.

[7] F. Saito and T. Suehiro, "Toward Telemanipulation via 2-D Interface - Concept and First Result of "Titi" -," *Proc. of IECON 02*, Vol.3, pp. 2243-2248, 2002.

[8] J. Neubert, P. Ravindran and N. J. Ferrier, "An Assistive Robot Interface for an Interactive Visually Guided Grasping System," *Proc. of the ICORR 2003*, pp. 290–293, 2003.

[9] K. Ikeuchi and T. Suehiro, "Towards an assembly plan from observation: Part 1: Assembly task recognition using face-contact relations (polyhedral objects)," *Proc. of the 1992 IEEE Int. Conf. on Robotics and Automation*, pp. 2171–2177, 1992.

[10] J. Miura, Y. Yano, K. Iwase and Y. Shirai, "Task Model-Based Interactive Teaching," *Proc. of IROS 2004 Workshop on Issues and Approaches to Task Level Control*, 2004.

[11] E. Ono, K. Nagata and Y. Wakita, "Indications for Handling of the Towel by Two-finger Parallel Gripper," *Proc. of ROBOMEC 2004*, (in Japanese) 2004.