Manipulation and Recognition of Objects Incorporating Joints by a Humanoid Robot for Daily Assistive Tasks

Mitsuharu Kojima, Kei Okada and Masayuki Inaba

Abstract—Methods for a daily assistive humanoid robot to manipulate and recognize the objects incorporating joints and learn the manipulation knowledge are presented. It is necessary for humanoid robots to use the objects incorporating joints such as some furniture and tools to provide daily assistance. We have been tried to make an integrated humanoid robots recognition and manipulation system of the objects and tools in the real world. We extend the system for the objects incorporating joints. In this paper, a recognition system in which the robots recognizes the objects incorporating joints by the visual 3D object recognition method with multi-cue integration using particle filter technique and a manipulation system of them are shown. The search areas of the joints are automatically generated based on the manipulation knowledge. We present three key techniques to recognize and manipulate the objects incorporating rotational and linear joints. 1) Knowledge Description for Manipulation and Recognition of these objects 2) Motion Planning Method to Manipulate them 3) Recognition Method of them Closely Related to the Manipulation Knowledge. Moreover, a method for a person to teach the handle, one of manipulation knowledge, visually to the robot is shown. Finally, a daily assistive task experiment in the real world using these elements is shown.

I. INTRODUCTION

Humanoid robots which assist human activities in daily life are expected. Many researches have been done to realize humanoid robots in a daily environment [1], [2]. In order to achieve these tasks in the real world, the ability for these robots to use some objects, tools and furniture which are designed for human is important. Moreover, it is desirable that these robots have the ability to learn, by themselves or from humans, the knowledge which is necessary to use these objects. Some researches are done to realize such learning ability [3], [4].

In this paper, we focus our attention on the recognition and manipulation of the objects incorporating joints such as refrigerators, rooms' doors, microwaves and closet. We have developed knowledge based humanoid system that integrates both motion planning system and visual object recognition system [5]. We extend the system to recognize and manipulate the objects incorporating joints and introduce learning ability of manipulation knowledge.

Section II describes our knowledge based humanoid robot system especially focused on objects incorporating joints. The method to recognize and manipulate the objects incorporating joints is shown. In section III, we show a daily assistive task experiment to use a refrigerator in the real world. In

M. Kojima, K. Okada, and M. Inaba are with the Graduate School of Information Science and Technology, The University of Tokyo, Engineering Building No. 2, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan kojima@jsk.t.u-tokyo.ac.jp
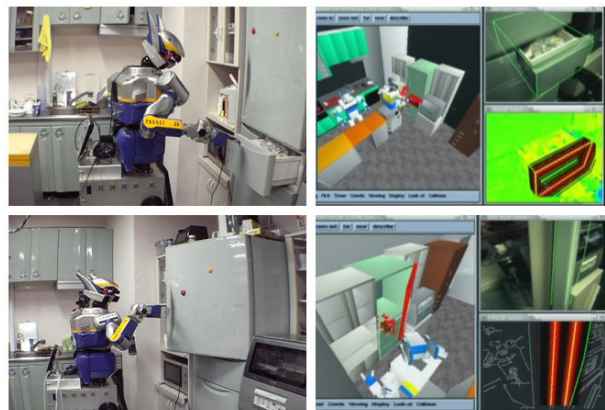

Fig. 1. Manipulation and recognition of the incorporated joints in a refrigerator (The robot manipulates the drawer in top left image, recognizes its linear joint angle in top right image, manipulates the door in bottom left image and recognizes its rotational joint angle in bottom right image, based on manipulation and visual feature knowledge.)

section IV, one of the learning abilities of manipulation knowledge is shown.

II. RECOGNITION AND MANIPULATION SYSTEM OF OBJECTS INCORPORATING JOINTS

Fig.1 shows the key-points of this paper. In top left image, the robot opens the refrigerator's drawer by the motion which is planned using the "*handle*" and "*linear joint*" knowledge. In bottom left image, the robot opens the refrigerator's door by the motion which is planned using the "*handle*" and "*rotational joint*" knowledge. In top right image, the robot recognizes the drawer's linear joint angle and in bottom right image, the robot recognizes the door's rotational joint angle. The search areas for the recognition are automatically generated from the "*joint*" knowledge. In the right images, the right bottom sub images show the joint recognition process. The red lines are super imposed drawer or door at the candidates' joint angles. In the right top sub images, the green lines indicate the super imposed drawer or door at estimated joint angle.

A. Knowledge about the objects incorporating joints

The knowledge such as "*handle*", "*joint*" is implemented on the humanoid robot programming system capable of three dimensional shape modeling [6], [7]. Fig.2 shows the implemented knowledge for a refrigerator.

- Manipulation Knowledge

– **Spot** is standard coordinates for tasks. When the robot is given a task, the robot moves globally to the "*spot*" for the task, then moves around the "*spot*" locally.

– **Handle** is target posture and position of the robot's end effector to manipulate the object. "*Handle*" is fixed to the object. So when the object moves, "*handle*" is also updated. Using "*handle*" knowledge, the motion planner generates motion in the following steps.

1) Move the object to the intended posture and position.
2) Update "*handle*" to follow the object motion.
3) Solve inverse kinematics to move the end effector to the "*handle*".

– **Joint** is movable mechanism of objects. "*Rotational and linear joint*" are implemented. For example, drawer has a "*linear joint*" and door has a "*rotational joint*".

- Visual Feature Knowledge [5]

  – **3D Edges** are one of the visual feature knowledge in the visual 3D object recognition method with multi-cue integration using particle filter technique [5], [8]. In the recognition system, likelihood between "*3D Edges*" and "*2D edge segments*" extracted from input image is calculated. This visual feature is effective for the recognition of *less textured* objects. "*Color histogram*" and "*3D shape*" ("*3D feature points*" in input image) are also implemented as visual feature knowledge. "*3D feature points*" are effective for *more textured* objects.

  – **Search Area** is previously defined or automatically generated search area for the object recognition.

These knowledge are shown in Fig.2. The red three lines are the "*spot*" for the task to use the refrigerator. The green triangles are "*handle*" of the drawers and the door. The blue cylinder is the "*rotational joint*" of the refrigerator's door. The white lines are "*3D edges*" of visual feature knowledge for self-localization. The red rectangle is a "*search area*" for self-localization.

B. Recognition of incorporated joints in objects

1) Likelihood calculation in the recognition system: In the recognition system, the following likelihood calculations are integrated.

- **I: Likelihood between 3D edges and 2D edge segments:** "*3D edges*" are visual feature knowledge and "*2D edge segments*" are extracted from input image. The shorter the distance and the angle between them are, the higher the likelihood is.

- **II: Likelihood between 3D shape and 3D feature points:** "*3D shape*" is visual feature knowledge and "*3D feature points*" are extracted from input image. "*3D feature points*" are extracted as followings.

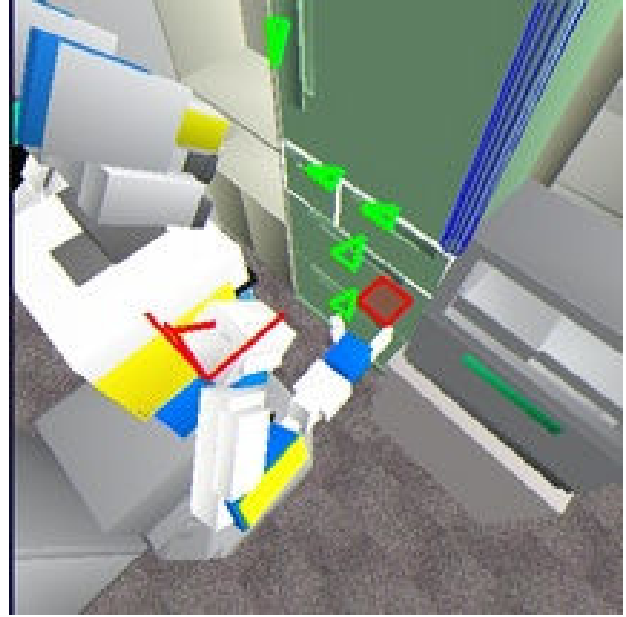 1) The 2D feature points are generated by using the KLT feature extraction method [9].

Fig. 2. Manipulation and visual feature knowledge of a refrigerator (Green triangles are "*handle*" and a blue cylinder is a "*rotational joint*" (Manipulation knowledge). White "*3D edges*" are visual knowledge for self-localization.)

2) The correlation based stereo matching is applied to calculate the disparities of the feature points.
3) The 3D distances of the points from the camera origin are calculated by assuming that the internal and external camera parameters are calibrated.

The shorter the distance between the faces of "*3D shape*" and "*3D feature points*" is, the higher the likelihood is.

- **III: Likelihood between color histograms:** the likelihood is calculated using the Bhattacharyya distance [10].

2) Search area generation for recognition of incorporated joint in objects, based on manipulation knowledge: We extend the recognition system to recognize the objects incorporating joints. The joint recognition closely relates its manipulation knowledge. The "*search area*" for recognition of the incorporated joints is automatically generated from the manipulation knowledge. For a rotational joint, a search area around the axis is generated, and its search angle range is determined by the joint max/min angles. For a linear joint, a search area whose search range is determined by the joint max/min positions is generated.

Fig.3 shows the generated search areas and visual feature knowledge. In left and right figures, search areas of linear joints of drawer are shown as red cylinders. In middle, a search area of a rotational joint of door is shown as a blue arc. In all three images, the red lines are the "*3D edges*" of visual feature knowledge for the joint recognition.

3) Joint recognition experiments: Fig.4 shows joint recognition experiments. In bottom row, the kitchen and refrig-
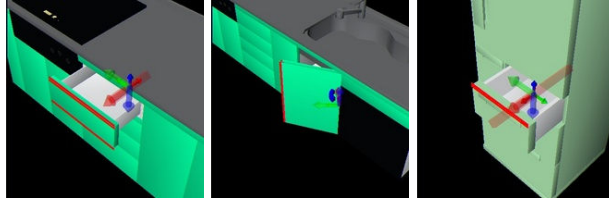
Fig. 3. Visual feature knowledge and "*search area*" for joint recognition (The red lines are "*3D edges*" of visual feature knowledge for a kitchen drawer and door and a refrigerator drawer. The red cylinders show the "*search areas*" of "*linear joints*" in right and left images. The blue arc shows the "*search area*" of the "*rotational joint*", in middle image.)
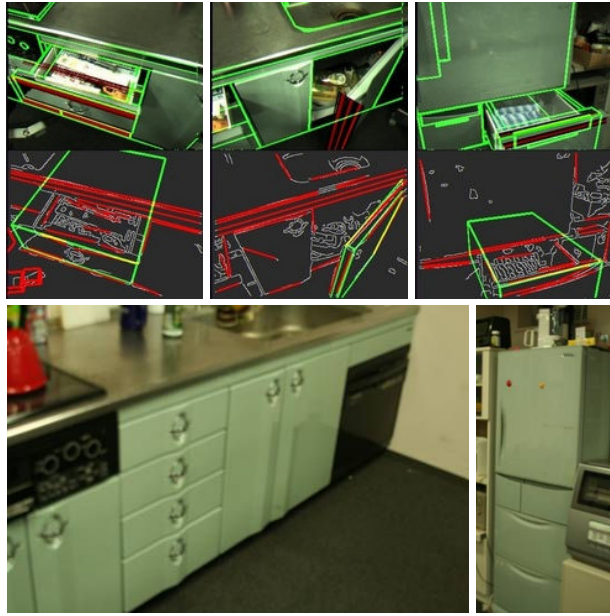

Fig. 4. Joint recognition experiments (Upper row images show joint recognition of a kitchen drawer, its door and a refrigerator's door in the lower row scenes.)

erator for the experiments are shown. In top row, joint recognition results are shown. The search areas and visual feature knowledge for the experiments are based on those of Fig.3.

In upper images of the top row images, the recognized models are super imposed in the green lines and the red lines show the candidates of joint recognition results. The redder the lines are, the higher the likelihood of the candidate is. The lower images are the edge segment extraction results and the green lines are the estimated results. Appropriate likelihood of the candidates and correct recognition results are shown.

C. Motion planning for the objects incorporating joints

Fig.5 shows the motion planning process to open the refrigerator's door. As described above, the geometric motion is planed in these steps, using knowledge of the objects incorporating joints.

1) Move the door by updating the "*rotational joint*" angle
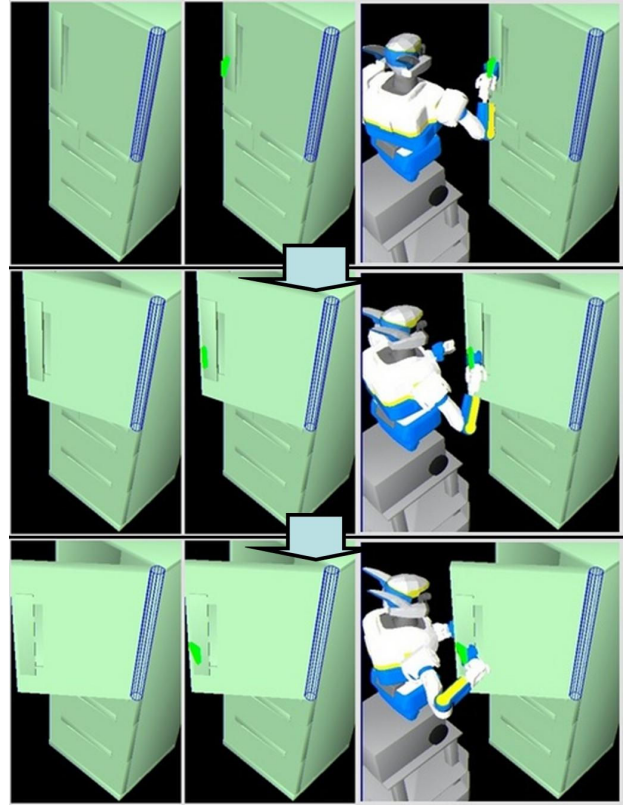

Fig. 5. Motion planning to manipulate the refrigerator's door (First column figures show the door movements by moving the "*rotational joint*". Second column figures show the "*handle*" movements by updating the "*handle*" to follow the "*joint*" motion. Third column figures show the robot movements which are solution of the inverse kinematics.)

(The first column figures of Fig.5). The blue cylinders are the "*rotational joint*".

2) Update "*handle*" to follow the door motion (The second column figures of Fig.5). The green triangles are the "*handle*" of the door.

3) Solve inverse kinematics to move the right arm's end effector to the "*handle*" (The third column figures of Fig.5). The robot posture changes, keeping the end effector grasp the "*handle*".

The motion plan result is applied to real robot with real-time impedance control of hands position.

Fig.6 is another example of motion planning to open the refrigerator's drawer. In left column, the planning results by the method are shown. In right column, the real motions are shown.

III. A DAILY ASSISTIVE TASK EXPERIMENT OF A REFRIGERATOR

A. Task scenario

Fig.7,Fig.8 and Fig.10 show the daily assistive task experiment. The task scenario is as followings.

1) Open the refrigerator's door (Fig.7). Self-localization and manipulation of the door incorporating a joint are necessary.
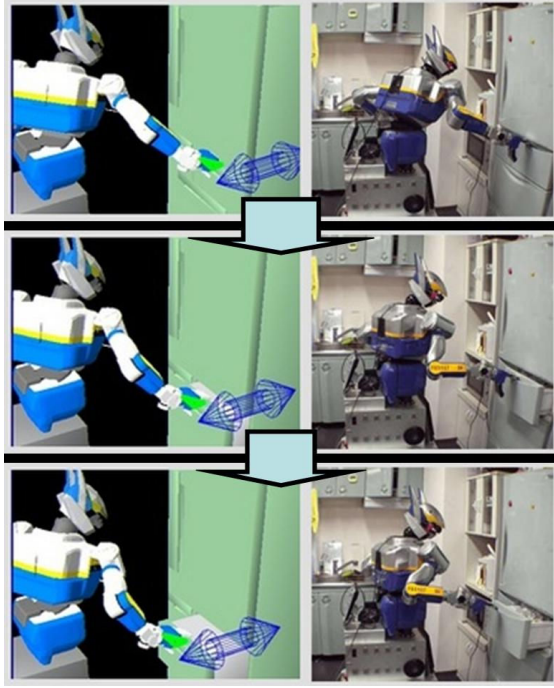
1566

Fig. 6. Motion planning to manipulate the refrigerator's drawer and the experiment on a real robot (Left column figures are planning results and right column figures are the real motions.)

2) Get out a plastic bottle from the refrigerator (Fig.8). Joint recognition of the door and object recognition of the plastic bottle in it are necessary.
3) Close the refrigerator's door (Fig.10). Joint recognition of the door and manipulation of the door are necessary.

B. Phase 1 : Open the refrigerator's door

Fig.7-1 shows the initial setting of the experiment.

1) Self-localization: First, the robot localizes its own position, based on object recognition of the refrigerator. Fig.7-2,3 show the image processing results of the self-localization. In the recognition, "*3D edges*" in Fig.2 are used as visual feature knowledge. The red lines in the lower images in Fig.7-2,3 show the super imposed "*3D edges*" at the candidates' positions of the refrigerator. In Fig.7-2, the position candidates are scattered and the estimated position has low reliability which is indicated by the blue super imposed lines. In Fig.7-3, the position candidates' have converged enough and the estimated position has high reliability which is indicated by the green super imposed lines.

2) Manipulation of the door incorporating a joint: Using the "*handle*" knowledge, the robot plans motion to open the door in the method described in Section II-C. Fig.7-4~11 are the planned motions. Fig.7-4,5 show the robot grasps the refrigerator's door from different view angles. Fig.7-6,7 shows the robot opens the door. Then the robot opens the door further in Fig.7-8~11.
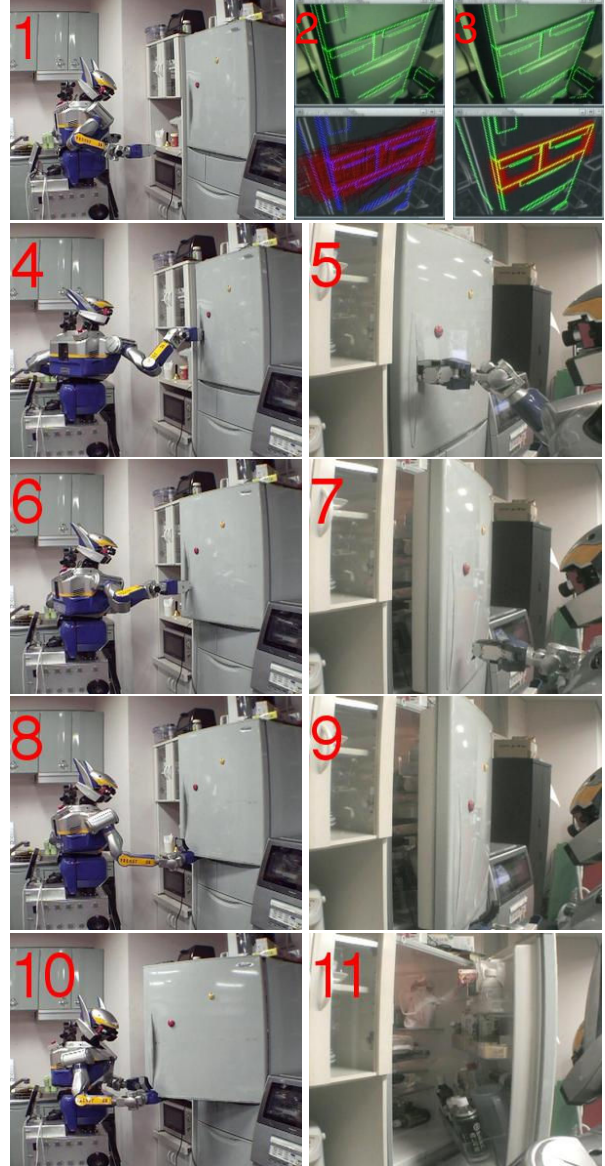

Fig. 7. Self-localization and manipulation of a refrigerator in a daily assistive task experiment (The robot localizes its own position (2,3), based on object recognition of the refrigerator and manipulates a door incorporating rotational joint of the refrigerator (4~11) , based on the manipulation knowledge.)

C. Phase 2 : Get out a plastic bottle from the refrigerator

Fig.8-1,4 show the pose of the robot to recognize the door and plastic bottle.

1) Joint recognition of the door: Fig.8-2,3 show the image processing results of the joint recognition of the door. In the recognition, "*3D edges*" which are shown as the red lines in Fig.9 are used as visual feature knowledge. The red lines in the lower images in Fig.8-2,3 show the super imposed "*3D edges*" at the candidates positions of the door. In Fig.8-2, the position candidates are scattered and in Fig.8-3, the position candidates have converged enough.
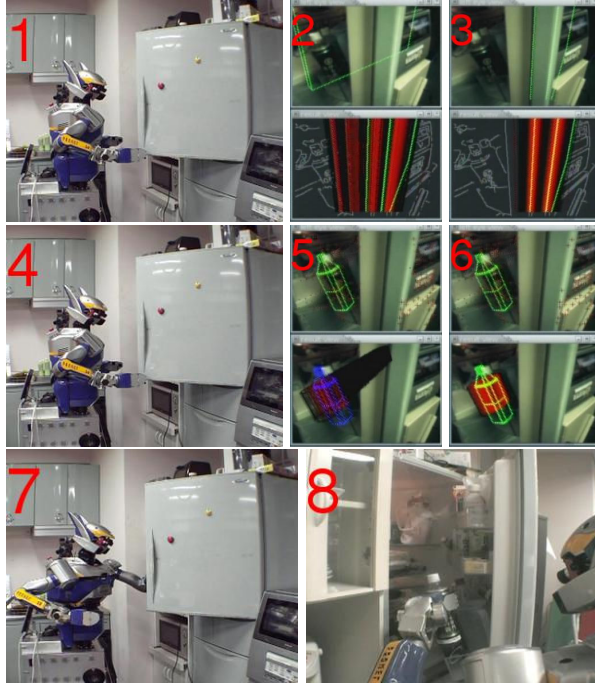
Fig. 8. Recognition of the rotational joint of the refrigerator (2,3) and object recognition (4∼6) and manipulation (7,8) of a plastic bottle in its.
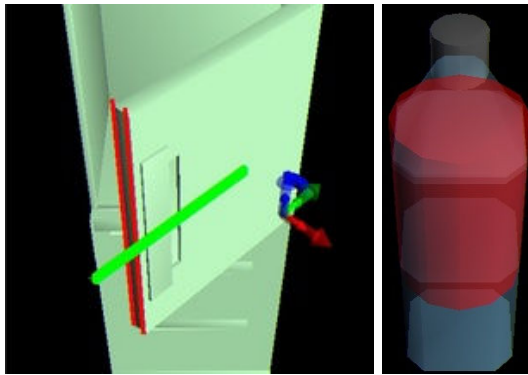

Fig. 9. Visual feature knowledge and "*search area*" (In left image, the red lines are "*3D edges*" of visual feature knowledge for joint recognition and the green region is search area of the plastic bottle. In right image, the red cylinder is the "*3D shape*" of visual feature knowledge.)

2) Object recognition of the plastic bottle: Fig.8-5,6 show the image processing results of the object recognition of the plastic bottle. Based on the joint recognition result, the search area of the bottle in the shelf on the door, which is shown as a green region in Fig.9, is determined. In the recognition, "*3D shape*" is used as visual feature knowledge which is shown in Fig.9. From Fig.8-5 to Fig.8-6, the position candidates become more convergent. In Fig.8-6, the estimated position has enough reliability.

3) Get out the plastic bottle: Based on the estimated plastic bottle's position, the robot gets out the bottle (Fig.8-7,8).
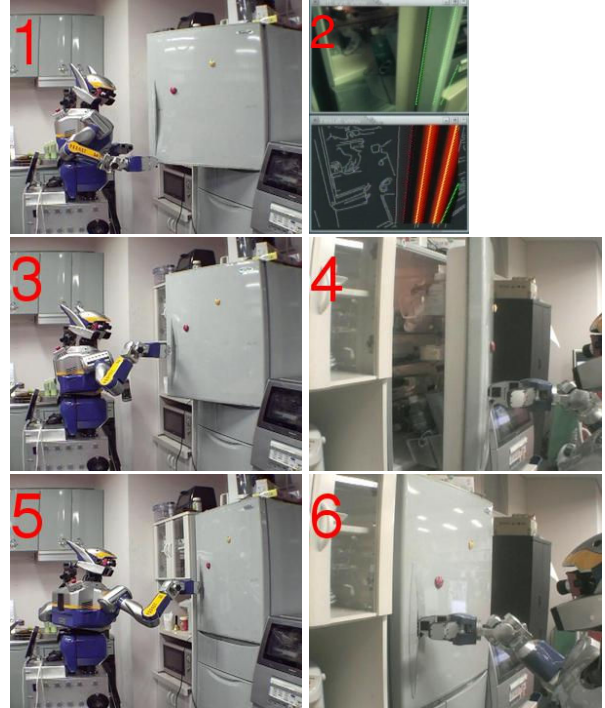

Fig. 10. Re-recognition of the rotational joint of the refrigerator (2) and manipulation of a refrigerator (3∼6)

D. Phase 3 : Close the refrigerator's door

1) Re-recognition of the door's joint angle: In the motions to get out the bottle, the door joint angle may have changed, so the re-recognition of the door's joint angle is important. Fig.10-2 shows the result of the re-recognition of the door's joint angle.

2) Close the refrigerator's door: Then the motion to close the door is planned, based on the recognized door's joint angle and using "*handle*" and "*joint*" knowledge by the method described in Section II-C. Fig.10-3∼6 show the planned motion. Fig.10-3,4 are the same scene from different angles, in which the robot closes the refrigerator's door. In Fig.10-5,6, the robot has closed it.

IV. LEARNING MANIPULATION KNOWLEDGE FROM HUMAN

A. Learning manipulation knowledge with recognition of the object and human

In order to learn manipulation knowledge of some objects from human, it is necessary to recognize the target object and human motion. In general, the process to learn manipulation knowledge is described as below.

1) Recognize the target object about which the robot learns manipulation knowledge.
2) Recognize human motion, focused on their hands' movements or gaze direction.
3) Interpret human motion relating the object. For example, if a human gazes at a point on the object, it is necessary for the robot to understand the point would
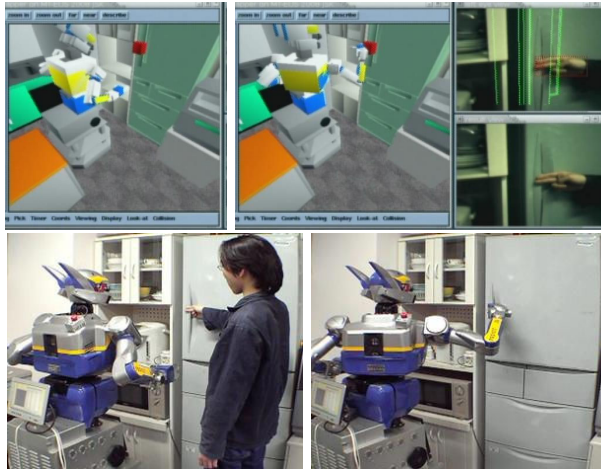
Fig. 11. Learning of a refrigerator's handle (A person visually teaches the robot the handle. In top left, the handle position which he teaches the robot is shown as a red cube (a). In top right upper, the green edges are super imposing of refrigerator position which is recognized previously and the red rectangle is a hand detection result (b). In top middle, the robot plans the motion to open the door (c). The left view of the bottom row is an external view for (a) and (b). The right view of the bottom row is the planed motion in (c).)

be an attention point of object at which the robot needs to gaze.

4) Manipulate the object, based on the learned manipulation knowledge and improve the knowledge with sensory information.

B. Learning manipulation of refrigerator

Fig.11 shows an experiment to learn manipulation knowledge of a refrigerator.

1) The robot has recognized the refrigerator's position. The green lines in top right image are super imposed refrigerator at recognized position.

2) A person teaches the robot the handle position by showing his hand. Its external view is shown in the left image of the bottom row. In the top right images, the robot extracts a human skin color region (a red rectangle), and then gets its 3D position by stereo block matching. The detected 3D position is shown as a red cube in top right and middle images.

3) If the detected 3D position is near to the refrigerator and unchanged for a while, the robot takes its position as a "*handle*" position.

4) The robot plans its motion to open the door based on the learned "*handle*", in the top middle image.

5) Finally, in the right bottom image, the robot tries to open the door and succeeds in it.

V. CONCLUSIONS

This paper presents recognition, manipulation and learning the manipulation of the objects incorporating joints.

The key points are:

1) Manipulation method of the objects incorporating joints.

2) Recognition method of the objects incorporating joints, closely relates their manipulation knowledge.

3) The recognition and manipulation method are checked in a daily assistive task experiment of a refrigerator.

4) To learn the manipulation knowledge, the recognition of object and human is necessary. As an example, a person teaches the robot a handle.

To realize humanoid robots which assist human activities in daily life, the robots must be able to manipulate and recognize the objects incorporating joints such as refrigerators and the learning ability of knowledge is necessary. In this paper, the manipulation and recognition system of such objects is described and the first step to the learning robots is shown.

The ability to learn more knowledge is required and the robots must be able to manipulate and recognize more various objects such as water, cloth and string. At present, if the room and objects are changed, the new 3D models should be rebuilt by human. Automatic 3D model construction is also necessary.

REFERENCES

[1] Y. Sakagami, R. Watanabe, C. Aoyama, S. Matsunaga, N. Higaki, and K. Fujimura. The intelligent ASIMO: system overview and integration. In *Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and System (IROS 2002)*, volume 3, pages 2478–2483, 2002.

[2] T. Asfour, K. Regenstein, P. Azad, J. Schroder, A. Bierbaum, N. Vahrenkamp, and R. Dillmann. ARMAR-III: An Integrated Humanoid Platform for Sensory-Motor Control. In *Proceedings of the 6th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2006)*, pages 169–175, 2006.

[3] D. Omrcen, A. Ude, K. Welke, T. Asfour, and R. Dillmann. Sensorimotor Processes for Learning Object Representations. In *Proceedings of the 7th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2007)*, 2007.

[4] O. Stasse, D. Larlus, B. Lagarde, A. Escande, F. Saidi, A. Kheddar, K. Yokoi, and F. Jurie. Towards Autonomous Object Reconstruction for Visual Search by the Humanoid Robot HRP-2. In *Proceedings of the 7th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2007)*, 2007.

[5] Kei Okada, Mitsuharu Kojima, Satoru Tokutsu, Toshiaki Maki, Yuto Mori, and Masayuki Inaba. Multi-cue 3D Object Recognition in Knowledge-based Vision-guided Humanoid Robot System. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, pages 3217–3222, 2007.

[6] Toshihiro Matsui. Multithread Object-Oriented Language EusLisp for Parallel and Asynchronous Programming in Robotics. In *Workshop on Concurrent Object-based Systems, IEEE 6th Symposium on Parallel and Distributed Processing*, 1994.

[7] Kei Okada, Takashi Ogura, Atushi Haneda, Daisuke Kousaka, Hiroyuki Nakai, Masayuki Inaba, and Hirochika Inoue. Integrated system software for HRP2 humanoid. In *Proceedings of the 2004 IEEE International Conference on Robotics and Automation (ICRA 2004)*, volume 4, pages 3207–3212, 2004.

[8] M. Isard and A. Blake. CONDENSATION – Conditional Density Propagation for Visual Tracking. *International Journal of Computer Vision*, 29(1):5–28, 1998.

[9] J. Shi and C. Tomasi. Good features to track. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 1994)*, pages 593–600, 1994.

[10] Thomas Kailath. The Divergence and Bhattacharyya Distance Measures in Signal Selection. *IEEE Transactions on Communications*, 15(1):52–60, 1967.