Proceedings of the 2002 IEEE/RSJ
Intl. Conference on Intelligent Robots and Systems
EPFL, Lausanne, Switzerland • October 2002

Teaching of Grasp/Graspless Manipulation for Industrial Robots by Human Demonstration

Yusuke MAEDA[1] Nanako ISHIDO[1] Haruka KIKUCHI[2] *and* Tamio ARAI[1]

[1]*Department of Precision Engineering, School of Engineering, The University of Tokyo, Tokyo, JAPAN, maeda@prince.pe.u-tokyo.ac.jp*

[2]*Mobile Agent Lab, Multimedia Labs, NTT DoCoMo, Inc., Yokosuka, JAPAN*

Abstract

In this paper, we propose a simple teaching method for industrial robots by human demonstration. The method can be divided into two parts: the teaching phase and the planning phase. In the teaching phase, a human operator demonstrates a manipulation of an object, and two cameras recognize the path of the object by observing markers attached on the object. In the planning phase, a motion planner generates a sequence of robot operations to play back the demonstrated path by pick-and-place and pushing. The proposed method incorporates automated camera calibration required for human demonstration, which enables labor-saving teaching and compensates the absolute positional error of industrial robots.

In the experiments, each of demonstrated manipulations was reproduced successfully as a combination of pick-and-place and pushing by one or two robot manipulators.

1 Introduction

The cost of robot teaching is one of the biggest issues that prevent the spread of robot utilization. Robot programming is not easy for novice operators and the cost of training them is often unaffordable especially for small companies. Thus low-cost and labor-saving robot teaching is greatly to be desired.

Automated robot motion planning by computer algorithms is a traditional approach for this problem. There exist a lot of algorithms for robot motion planning [1, 2]. However, they are uncommonly used in industrial applications. That is because additional procedures are necessary for the execution of planned motion; most of motion planning algorithms require a detailed model of the workspace of robots, and absolute position of real robots must be calibrated before the execution.

In recent years, the "teaching-by-showing" (or "programming-by-demonstration") approach has been widely studied. Some researchers adopted intensive image-processing to acquire task knowledge from human demonstration. For example, simple assembly tasks [3] and pick-and-place [4] were achieved. Ogata and Takahashi constructed a virtual environment where a human operator demonstrates assembly operations with virtual reality tools such as a data glove [5]. The demonstration is interpreted and replicated by a robot in a real environment. 3D position sensors are also used in some robot programming systems by human demonstration [6, 7]. Ogawara et al. used intensive image-processing and data gloves to analyze human demonstration [8].

Though these achievements are very important, the "teaching-by-showing" approach is still far from industrial applications. There are various reasons:

- Preparation for human demonstration is a hard task; much work such as modeling of the object and the environment is required even if teaching itself is easy.
- Special facilities for human demonstration (e.g. data gloves, position sensors, etc.) are expensive and/or of insufficient accuracy.
- Calibration of real robots is still necessary.

In this paper, we develop a simple "teaching-by-showing" method for grasp/graspless manipulation by conventional industrial robots; that is, pick-and-place and pushing operation [9] of robot manipulators is generated from human demonstration. We attach markers on manipulated objects to make it easy to recognize human demonstration. The paths of the objects are observed by two cameras and reproduced using pick-and-place and pushing by robots (Figure 1). We also attach markers on the end-effectors of the robots to realize automated calibration required for human demonstration.

2 Outline of Teaching

In this paper, we set the following assumptions:

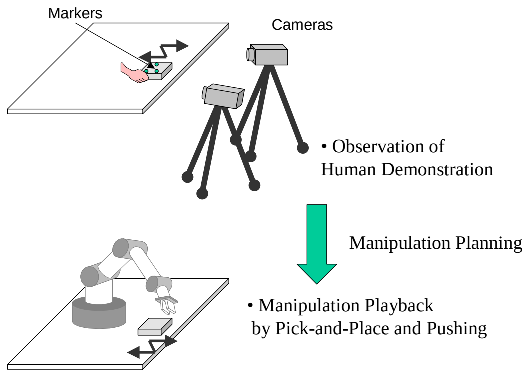


Figure 1: *Demonstration and Playback of Manipulation*

- We use conventional position-controlled robots equipped with parallel-jaw grippers.
- Each robot can perform pick-and-place and pushing operation of an object with its gripper.
- Manipulation tasks are executed on and above a flat table. Information on the workspace (e.g., obstacles) is given indirectly by human demonstration.
- The shape of the manipulated object and positions of markers on it in the object coordinate system are known.

Our teaching method consists of the following two phases:

1. Teaching Phase
   - Camera calibration for human demonstration using robot coordinates
   - Teaching of an object path by human demonstration
2. Planning Phase
   - Segmentation of the object path
   - Planning of robot motion for the path replication

We describe our method in detail in the following sections.

3 Teaching of Manipulation

3.1 Teaching of Object Path by Human Demonstration

In our method, a human operator manipulates an object from an initial position to a goal position. The path of the object (position and orientation) in the demonstration is used as teaching data. We adopt the DLT (Direct Linear Transformation) method [10] for obtaining 3D position data of markers attached to the object; position/orientation data of the object are calculated from the markers' positions.

The DLT equation is:

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} \frac{b_{i,1}x+b_{i,2}y+b_{i,3}z+b_{i,4}}{b_{i,5}x+b_{i,6}y+b_{i,7}z+1} \\ \frac{b_{i,8}x+b_{i,9}y+b_{i,10}z+b_{i,11}}{b_{i,5}x+b_{i,6}y+b_{i,7}z+1} \end{bmatrix}, \tag{1}$$

where $[u_i, v_i]^T$ is the position of a marker in the image plane of i-th camera ($i = 1, 2$), $\boldsymbol{x} = [x, y, z]^T$ is the 3D position of the marker, and $b_{i,j}$ are unknown constants (DLT parameters).

We rewrite eq. (1) as

$$\boldsymbol{u} = \boldsymbol{f}(\boldsymbol{x}), \tag{2}$$

where $\boldsymbol{u} = [u_1, v_1, u_2, v_2]^T$. Once all the DLT parameters ($b_{i,j}$) are identified by calibration, we can perform 3D reconstruction as:

$$\boldsymbol{x} = \boldsymbol{f}^{-1}(\boldsymbol{u}) \tag{3}$$

by the linear least-squares method.

Observation of three markers attached on the object enables us to obtain position/orientation of the object.

3.2 Calibration for Human Demonstration

The DLT parameters $b_{i,j}$ must be identified through calibration before 3D reconstruction by eq. (3). Easy calibration is indispensable to easy robot teaching; therefore we identify the DLT parameters automatically by observing a marker attached on the end-effector of each robot. (Figure 2).

Each robot can calculate the position of the marker from its own internal sensors (encoders). Thus we can obtain 3D position of the marker in the reference frame of the robot, $\boldsymbol{x}$, and its camera coordinates, $\boldsymbol{u}$. By moving the robot (typically, to point at an $n \times n \times n$ grid), we have multiple pairs $(\boldsymbol{x}_k, \boldsymbol{u}_k)$. The DLT parameters can be calculated by the linear least-squares method from $(\boldsymbol{x}_k, \boldsymbol{u}_k)$ to minimize the following residual:

$$\sum_k \|\boldsymbol{x}_k - \boldsymbol{f}^{-1}(\boldsymbol{u}_k)\|^2. \tag{4}$$

After the calibration, we can obtain the path of the object in human demonstration represented in the reference frame of the robot.

Note the following features of our calibration method:

- Positions of the cameras can be unknown, thus a human operator can place them with no special care just before the manipulation demonstration.

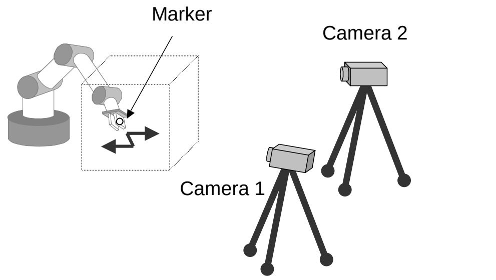


Figure 2: *Calibration for Human Demonstration*

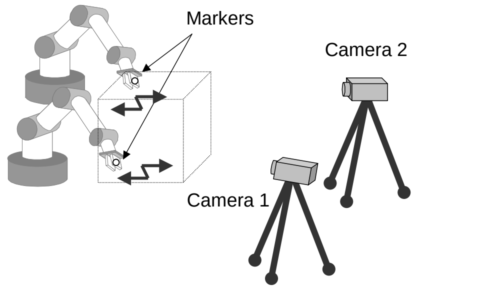


Figure 3: *Calibration for Human Demonstration (in a Multi-Robot System)*

- Robot motion required for the calibration can be automated, therefore all the human operator has to do is to place the cameras where they can observe the markers.
- Calibration is achieved using robot coordinates and consequently absolute positional error of the robots can be canceled.
- As we increase the number of the measurement points of the marker, the calibration accuracy is improved (and the calibration time is prolonged).

Calibration for multi-robot systems can be performed in a similar way (Figure 3). Homogeneous transformation matrices between robot coordinate systems are obtained in addition to the DLT parameters. Detailed description is found in [11].

4 Planning of Robot Motion for Manipulation Playback

After a path of the object is given in the teaching phase, robot motion should be generated to play back the path. First, we segment the path as a sequence of constrained and unconstrained motions. Then a planning algorithm generates robots' motion to reproduce each of segmented motions.

4.1 Segmentation of Object Path

In this research, we use conventional position-controlled robots. That is, the robots cannot perform compliant motions in contact tasks. Therefore, constrained contact motions of the object in demonstration have to be reproduced by pushing—in other words, we deal with only contact motions that can be performed by pushing. On the other hand, unconstrained motions can be reproduced by pick-and-place. Thus we segment the path of the object into constrained/unconstrained motions.

To detect constrained motions in the demonstrated path, we adopt the 3D Hough transformation. If the Hough transformation finds that a sequence of the path data lies on a plane, we identify the plane exactly by a least-squares method. If the residual is smaller than a threshold, we consider the sequence as a constrained motion.

4.2 Motion Generation for Manipulation

After the segmentation of the demonstrated path, we run a robot motion planner to reproduce each of the segments. We implemented a naive motion planning algorithm, which is sufficient for our experiments. There exist, however, a lot of elaborate motion planners (for example, [12]). Our motion planner can be replaced by a different one if necessary.

Planning of the following robot motions is required here:

1. Simple operation
   - Planning of pick-and-place
   - Planning of pushing
2. Transition of operation
   - Planning of transit motion from pick-and-place to pushing
   - Planning of transit motion from pushing to pick-and-place
3. Subdivision of operation
   - Planning of regrasping in single operation

Moreover, planning of operation assignment to a proper robot is necessary for multi-robot systems. Hereafter, each of them are described briefly.

Planning of Simple Operation. Basically, each segment of unconstrained motions is simply played back by pick-and-place operation. A robot grasps the object and moves it to the end of the segment.

On the other hand, each segment of constrained motions is usually reproduced by pushing operation. The robot approaches to the object from behind and pushes it to the end of the segment with its gripper.

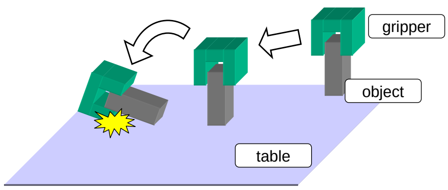


Figure 4: *Need for Regrasping in Single Operation*

Planning of Transition of Operation. The robot has to place its gripper to a suitable position for the next operation at a switching point between constrained/unconstrained motions. Such transit motions must be generated by the motion planner. In the case of switching from pick-and-place to pushing, the robot ungrasps the object and moves its gripper behind the object. On the other hand, in the case of switching from pushing to pick-and-place, the robot approaches to the object and grasps it.

Planning of Regrasping in Single Operation. The robot may have to regrasp the object in the middle of a single operation. For example, its gripper may hit against the table in a sequence of pick-and-place but for regrasping (Figure 4). In such cases, regrasping motion should be generated in an unconstrained motion by the motion planner.

Planning of Operation Assignment. In a process of motion planning, a planned path of the robot gripper is discretized into a series of points, where joint angles for a robot are calculated by inverse kinematics. If the joint angles are infeasible, the robot cannot complete the operation because of the limitation of its movable area. We check whether another robot can take over the operation when we use multiple robots. Thus the rest of the operation is assigned to a proper robot, if any.

Integration of all of the motion planning procedures described above enables us to generate robots' motion required to play back the demonstrated manipulation.

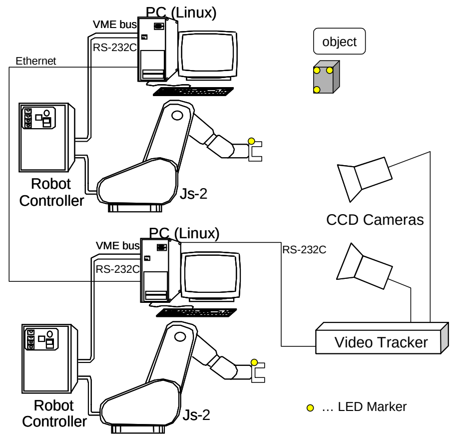


Figure 5: *Experimental Setup*

Figure 6: *Human Demonstration in an Experiment*

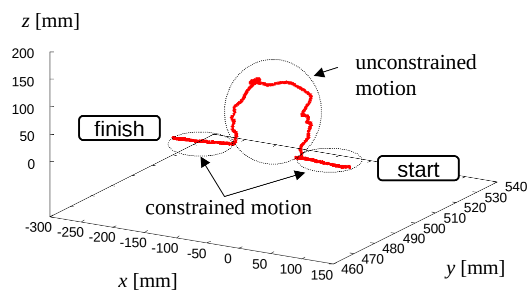


Figure 7: *Demonstrated Path in an Experiment*

5 Manipulation Experiments

5.1 Experimental Setup

Our experimental setup is illustrated in Figure 5. We use two 6-degree-of-freedom robots, "Js-2," which are position-controlled at 16 [ms] intervals by Linux PCs. We attach an LED marker on each hand of the robots. For the DLT method, a 3D motion measurement system are installed. It consists of two CCD cameras with 640×416 resolution and an arithmetic unit for the DLT, "VideoTracker G280" by OKK Inc. In our experiments, the accuracy of 3D reconstruction by the DLT is about 1[mm]. Note that we can replace the arithmetic unit by a PC equipped with a image processing board. We use cuboid-shaped cartons as manipulated objects. Three LED markers are attached on them.

5.2 Manipulation by a Single Robot

Manipulation consists of constrained and unconstrained motions was demonstrated (Figure 6). In this case, the demonstrated path of the object was

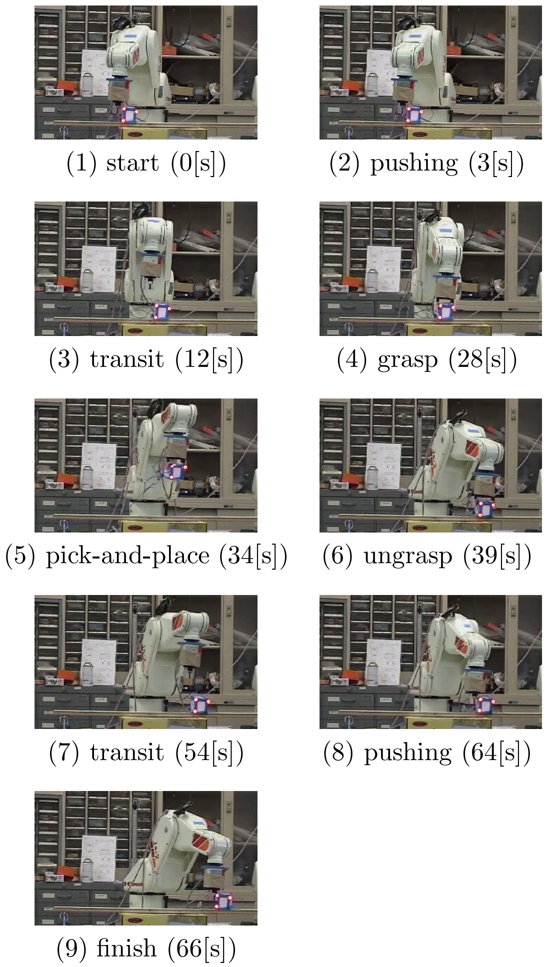


Figure 8: *Manipulation Playback by Pick-and-Place and Pushing*

segmented into three parts: two constrained motions and one unconstrained motion (Figure 7). The path was successfully played back by a single robot, using pick-and-place and pushing (Figure 8).

We demonstrated another manipulation consists of a single pick-and-place operation. In this case, our motion planner detected the danger of collision of the robot hand against the table. Therefore the robot played back the path of the object by pick-and-place with regrasping (Figure 9).

In the first case, the total time required for robot teaching was about 100[s]: 60[s] for calibration (with eight measurement points), 20[s] for human demonstration, and 20[s] for path segmentation (by a PC with Pentium III–600MHz). Time for motion planning was negligible.

5.3 Manipulation by Dual Robots

Another manipulation consists of constrained and unconstrained motions was demonstrated for a dual

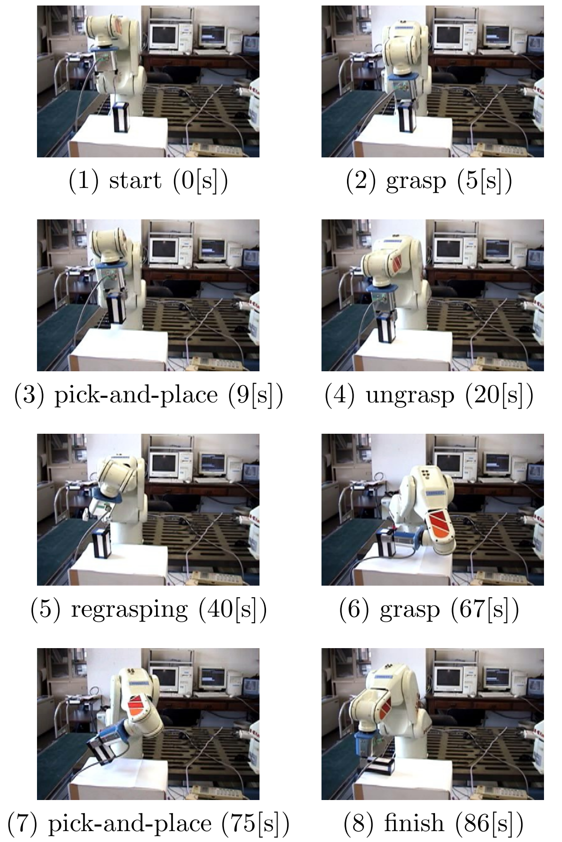


Figure 9: *Manipulation Playback with Regrasping*

robot system. The demonstrated path of the object was also successfully played back by two robots, using pick-and-place and pushing (Figure 10).

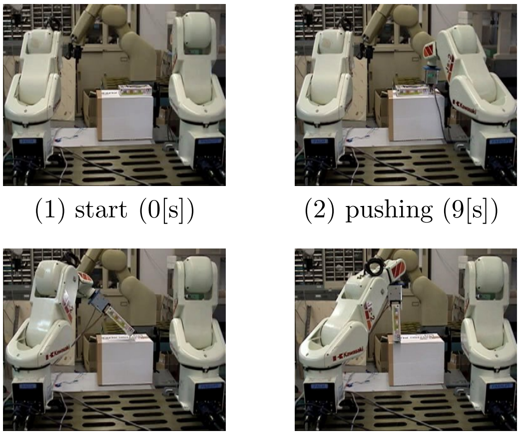
In this case, the total time required for robot teaching was about 160[s]: 120[s] for calibration (with eight measurement points for each robot), 20[s] for human demonstration, and 20[s] for path segmentation.

6 Conclusion

In this paper, we developed a simple teaching method for industrial robots by human demonstration. The method consists of two parts: the teaching phase and the planning phase. In the teaching phase, a human operator demonstrates a manipulation of an object, and the path of the object is obtained by observing markers attached on the object with two cameras. In the planning phase, a sequence of robot motions is generated to reproduce the demonstrated path using pick-and-place and pushing.

Features of our teaching method can be summarized as follows:

- Use of markers reduces total cost of robot teach-



(1) start (0[s]) (2) pushing (9[s])

(3) pick-and-place (148[s]) (4) finish (154[s])

Figure 10: *Manipulation Playback by Dual Robots*

ing including the labor required to set up the teaching system.

- Calibration for human demonstration is incorporated in the method and most of the calibration procedures can be automated. After the calibration, robots can recognize demonstrated paths of objects in their own reference frames. That means absolute positional error of the robots can be canceled to a certain extent.
- Special knowledge on the robot system (for example, movable area of each robot) is not very necessary for human demonstrators. Even if the system consists of multiple robots, the demonstrators only have to show the desired path of the object.

In the experiments, demonstrated manipulations were successfully reproduced using pick-and-place and pushing operation by robot manipulator(s).

The proposed teaching method can be used to support "Plug & Produce" [13], which is a function of manufacturing systems to realize their reconfigurability. Low-cost robot teaching will enhance the reconfigurability of manufacturing systems, because robots can be easily installed and joined in work immediately.

Occlusion of the markers in human demonstration is the biggest limitation in our current implementation. Future work should address the use of four or more markers and the estimation of positions of occluded markers.

Acknowledgments

This research is supported by IMS/HMS (Intelligent Manufacturing Systems/Holonic Manufacturing Systems) International Research Program.

References

[1] J.-C. Latombe: "Robot Motion Planning," Kluwer Academic Publishers, 1991.

[2] J.-C. Latombe: "Motion Planning: A Journey of Robots, Molecules, Digital Actors, and Other Artifacts," Int. J. of Robotics Research, Vol. 18, No. 11, pp. 1119–1128, 1999.

[3] K. Ikeuchi and T. Suehiro: "Toward an Assembly Plan from Observation, Part I: Task Recognition with Polyhedral Objects," IEEE Trans. on Robotics and Automation, Vol. 10, No. 3, pp. 368–385, 1994.

[4] Y. Kuniyoshi, M. Inaba and H. Inoue: "Learning by Watching: Extracting Reusable Task Knowledge from Visual Observation of Human Performance," IEEE Trans. on Robotics and Automation, Vol. 10, No. 6, pp. 799–822, 1994.

[5] H. Ogata and T. Takahashi: "Robotic Assembly Operation Teaching in a Virtual Environment," IEEE Trans. on Robotics and Automation, Vol. 10, No. 3, pp. 391–399, 1994.

[6] N. Delson and H. West: "Robot Programming by Human Demonstration: The Use of Human Variation in Identifying Obstacle Free Trajectories," Proc. of 1994 IEEE Int. Conf. on Robotics and Automation, pp. 564–571, 1994.

[7] H. Friedrich, J. Holle and R. Dillmann: "Interactive Generation of Flexible Robot Programming," Proc. of 1998 IEEE Int. Conf. on Robotics and Automation, pp. 538–543, 1998.

[8] K. Ogawara, S. Iba, T. Tanuki, H. Kimura and K. Ikeuchi: "Recognition of Human Task by Attention Point Analysis," Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp. 2121–2126, 2000.

[9] M. T. Mason: "Mechanics and Planning of Manipulator Pushing Operations," Int. J. of Robotics Research, Vol. 5, No. 3, pp. 53–71, 1986.

[10] R. Shapiro: "Direct Linear Transformation Method for Three-Dimensional Cinematography," Research Quarterly, Vol. 49, No. 2, pp. 197–205, 1978.

[11] T. Arai, Y. Maeda, H. Kikuchi and M. Sugi: "Automated Calibration of Robot Coordinates for Reconfigurable Assembly Systems," Annals of the CIRP, Vol. 51, No. 1, 2002 (to appear).

[12] H. Terasaki and T. Hasegawa: "Motion Planning of Intelligent Manipulation by a Parallel Two-Fingered Gripper Equipped with a Simple Rotating Mechanism," IEEE Trans. on Robotics and Automation, Vol. 14, No. 2, pp. 207–219, 1998.

[13] M. Sugi, Y. Maeda, Y. Aiyama and T. Arai: "Holonic Robot System: A Flexible Assembly System with High Reconfigurability," Proc. of 2001 IEEE Int. Conf. on Robotics and Automation, pp. 799–805, 2001.