8-12

MVA2007 IAPR Conference on Machine Vision Applications, May 16-18, 2007, Tokyo, JAPAN

Furniture Model Creation based on Direct Teaching to a Mobile Robot

Kimitoshi Yamazaki
Grad. School of Syst. and Inf. Eng.
Univ. of Tsukuba
Tsukuba, Japan
yamazaki@roboken.esys.tsukuba.ac.jp

Takashi Tsubouchi
Grad. School of Syst. and Inf. Eng.
Univ. of Tsukuba
Tsukuba, Japan
tsubo@roboken.esys.tsukuba.ac.jp

Masahiro Tomono
Dept. of Syst. Robotics
Toyo Univ.
Saitama, Japan
tomono@eng.toyo.ac.jp

***Abstract*— In this paper, a modeling method to handle furniture is proposed. In real environment, there is a lot of furniture such as drawer, cabinet and so on. If mobile robots can handle these furniture autonomously, it is expected that multiple daily tasks, for example, storing a small object in a drawer, can be performed by the robots. However, it is a perplexing process to give several pieces of knowledge about the furniture to the robots manually. So we propose direct teaching based approach which can easily give not only how to handle the furniture but also an appearance and 3D shape of it simultaneously. The main topic is about the appearance and shape information acquisition using external sensors.**

I. Introduction

Recently, various robots such as humanoids are actively developed and these robots are expected to work in behalf of human in near future. Because these robots have both human-like many joints and moving ability in real environment, it is desired to make the robots doing our daily tasks. For example, carry an object and store it in drawers, cabinets and so on.

In several researches, robot systems were developed to be assumed to do such tasks in real environment. There were specific requirements as follows: (1) almost knowledge for handling is given manually [4] [5], or (2) ID tags or marks are given to objects [2] [3], or (3) handling procedure is given by tele-operation [6]. Our emphasis is that these approaches cause to come on the human when they want to increase handleable objects for the robots. So it is desirable to reduce these manual preparation because there are a lot of objects which can be handled by the robot in real environment.

The purpose of this research is to create furniture model which enables robots to handle it. We propose direct teaching based method that human can easily give such model to the robot. Furniture model proposed in this paper is represented as composite model including not only how to handle it but also its appearance and 3D shape. We call this model ”IM-model (Instructed Motion Model)”. By using this model, it is expected that we can add handleable targets to the robot without specific tools or detail instructions.

For creating the IM-model, we need some external sensors which are mounted on a mobile manipulator. In this paper, a single camera and SOKUIKI sensor [7] (compact Laser Range Scanner) are utilized. We propose efficient model creation method by utilizing these sensor and direct teaching. The main topic in this paper is acquiring an appearance and shape information of furniture by using the camera and the SOKUIKI sensor.

II. Issues and Approaches

A. Assumed Task

As an easily assumed daily task, we set ”a robot conveys a small object and stores it in known furniture”. Furniture is assumed that it is needed to handle for storing task, for example, drawers, cabinets or sliding doors. In a drawer case, the robot has to draw it for putting a small object in.

B. Approaches

We aim to develop an easy extensional modeling method for the robot which can handle furniture. Our method is based on direct teaching [1] which needs no specific tools and detail instructions. We join this teaching with external sensing for acquiring appearance and 3D shape information of the furniture by using a camera and a SOKUIKI sensor.

The challenges of our approach are as follows: (1) how to extract modeled information from manually instructed data, (2) how to extract needed information from external sensor data. We solve these along following manners.

Extract handling model from manually instructed data: Our particular assumption is that almost furniture in real environment can classify as shown in TABLE I relying on its handling way. Using this assumption, we give the robot generalized hand motion obeying each classification. Because there are few classes, it is needed less manually efforts for defining these knowledge. In other words, an advantage of our approach is that one knowledge can be adopted to several types of furniture in common.

TABLE I

Variation of Furniture Manipulation

Type	Manipulation	Example
Drawer	pulling horizontal back or front trajectory	Cabinet
Sliding Door	pulling horizontal right or left side trajectory	Cupboard
Rotating Door	pulling rotational trajectory along fixed vertical axis	Refrigerator

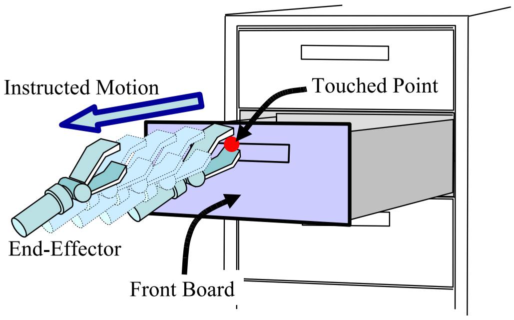


Fig. 1. IM model

On the other hand, because each of the furniture has various size or appearance, direct teaching is adopted for modeling in each furniture. However, all of acquired data from direct teaching is not always accurate, each model is modified by utilizing classified knowledge. Detail explanation is left out this paper.

Extract shape and appearance from external sensor data: It is assumed that a single camera and a SOKUIKI sensor are mounted on the robot. In our policy, the 3D shape and the appearance of furniture are modeled using image data and scanned data. This is created as front board model including 3D contours and color pixels inner contour. The former is acquired from projected image edges to scanned data. The latter is acquired by cutting from image. Detail explanations are described in the next Section.

These information are integrated to our model representation. We call this model "**IM-model** (Instructed Motion Model)" because it is created through manual instruction.

III. Modeling Outline

In this section, IM-model creation method is outlined. One of the IM-model component is front board information as appearance and 3D shape. Front board (Fig.1) is a part of furniture where its handled motion reflects the instruction directly.

A. Teaching Procedure

Teaching is performed as following procedure:

1) after installing a robot near to furniture, teaching is started. The robot measures the furniture by external sensors for creating initial model,
2) direct teaching is performed, that is, human handles a robot arm and a mobile platform directly for instructing how to handle the furniture.
3) after teaching, one image is captured. Finally, IM-model is created using these sensor data and handled motion of the robot.

Now, above items are explained in detail.

B. Initial Model Creation

Initial model means 3D contour candidates of a front board. This is acquired before teaching through following procedure. First, using the SOKUIKI sensor, dense 3D points are measured with related to environment in front of the robot (in our assumption, the scanner type sensor will be extended to 3D space scan not only 2D plane). Moreover, triangle patches are adopted to these 3D points. On the other hand, 2D line segments are extracted from image edges (Fig.2,(1)left) which are extracted from firstly captured image. 2D contour candidates are constructed from these line segments by combining these segments to quadrangle (Fig.2,(2)left). Finally, these 2D quadrangles are projected to 3D triangle patches, 3D quadrangles are acquired as contour candidates of the front board.

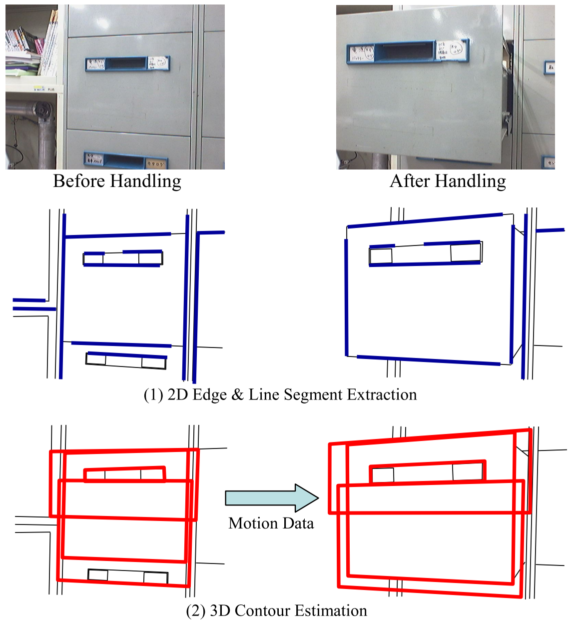


Fig. 2. Front board estimation

C. Trajectory Specialization

During the teaching, the robot records its directly handled motion by periodic interval. Because a part of this data indicates the information of the furniture handling, such part has to be extracted from all of the teaching data. Fig.3 shows an example of the trajectory for drawing motion. At first (1) Human leads the robot hand directly and starts the teaching from initial pose of the robot. Next, (2) the robot hand is inserted to the knob of the drawer, and (3) handling the drawer. After that, (4) releasing the hand, and (5) putting the manipulator back to initial pose. In this flow, (2) to (4) are needed for creating IM-model. So posterior knowledge is given to the robot for extracting these information and modeling this handling procedure.

D. Front board selection from its candidates

The contour candidates of the front board can include several impertinent ones which are extracted from around environment. So it is needed to select proper contours by using image data which is captured after teaching.

The procedure is as follows: at first, after the robot hand is returned to initial pose, an image is captured. Because handled motion of the furniture is already known

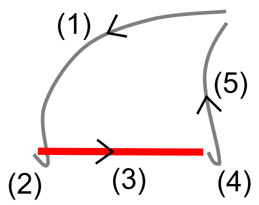


Fig. 3. Trajectory Hand in drawer case

through teaching, the present state of the front board is estimated as shown in (Fig.2,(2)right). By superimposing these estimated result on the image edges which is captured after teaching (Fig.2,(1)right), well matched contours are selected. This is described in the next section in detail.

According to these process, final IM-model is acquired as an approximated front board model which consists of several 3D contours, image data and handled information. By registering this model in database, it is expected to the robot automatic task implementation with finding the furniture autonomously because of its recognition using the model.

IV. Front Board Modeling based on Image Edges

In this section, the contours extraction of a frond board is described.

A. Approach

In initial modeling described in III.B, many 3D contours are created as front board candidates. These consist of 3D line segments which construct quadrangle.

To select true contours which represent the front board, we take an approach to project 3D line segments on image for corresponding to image edges. Because several candidates do not belong to the front board as described in III.D, image edges are utilized to confirm that the 3D line segments indicate really contours of the front board.

There is two challenges for this confirmation.

1) **Correspond the candidates with image edges**
In the case of our research, because only a part of environment is moved through teaching, it is difficult to correspond image edges which is captured after handling with the predicted contours.
Although feature points are often used to extract characteristic information from image robustly, this method is useless in our case because there is almost no correspondence on the boundary between the handled part and surroundings.
2) **Error of camera pose and 3D shape**
Basically, camera poses are estimated by such motion data of the robot as odometry or joint angles. This may have measurement error. In addition, line segments which are the material of 3D contour reconstruction may have measurement error, too. It is impossible to cancel these error completely.

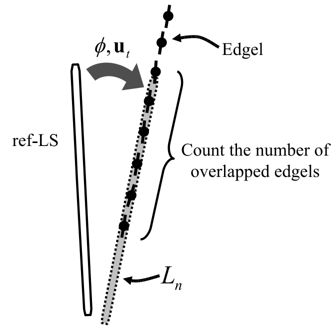


Fig. 4. Line segment vs. edgels

We propose a contour corresponding method based on image edges. The correspondence is judged by evaluating how many image edgels overlaps with a projected line segment of contour. The problem is that this matching will not be completed because of the environment changing and the measurement errors.

So, we propose two step matching as follows: at first, image level matching is performed that several image edges are corresponded to one line segment of the contour. Next, using the relation of perspective projection between camera pose and 3D line segment, the correspondences are evaluated. As a result, line segments which have high reliability are selected.

B. Correspondence based on line segment vs. edgels

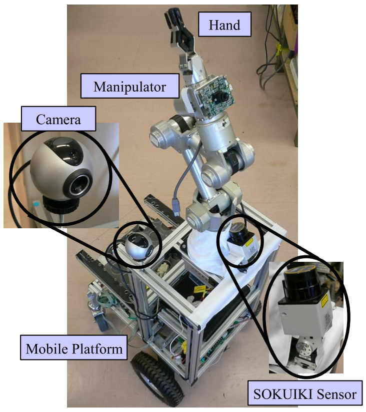
In the initial modeling, contour candidates are reconstructed as 3D line segments, and through direct teaching, motion of the front board can be measured. These information enable to estimate how the contour candidates are projected in image which is captured after teaching.

Now, we call the line segments of projected contour "ref-LS". The aim of the process is to evaluate the matching between ref-LS and image edges. To extract image edges, canny operator is utilized for calculating differential image, and thin edges are extracted by tracking a peak intensity. Next, as shown in Fig.4, slightly rotation ϕ and translation $\mathbf{u}_t$ are added to the ref-LS, correspondence is evaluated between the L_n (slightly translated ref-LS) and image edgels. If L_n and image edgels overlap with high ratio, the L_n (ref-LS, ϕ and $\mathbf{u}_t$) is registered. In this process, the correspondence will not be always completed, several correspondences are permitted to register against one ref-LS. In our implementation, the maximum number of n is set as 10.

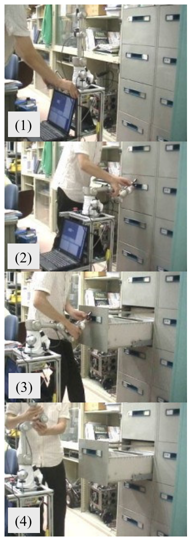
C. Reliability evaluation of correspondence

The correspondence process described in former subsection is performed 2D image coordinates. To verify the correspondence consistency, L_n are evaluated in 3D coordinates. Evaluation function is defined as follows:

$$F = \sum_{i=1}^{m} \left(\rho_n - f \frac{\{\cos\theta_n(\mathbf{r}_1\mathbf{x} + t_1) + \sin\theta_n(\mathbf{r}_2\mathbf{x} + t_2)\}}{\mathbf{r}_3\mathbf{x} + t_3} \right), \quad (1)$$



A Mobile Manipulator



Teaching Procedure

Fig. 5. Experiment

where ρ_n and θ_n indicate hough parameters of L_n. $\mathbf{x}$ indicates the 3D coordinates of the center of each line segments (3D ref-LS before projected to image). $\mathbf{r}_i$ and $t_i (i = 1, 2, 3)$ indicate camera rotation and translation parameter. f is focal length. m is the number of L_n which is randomly selected.

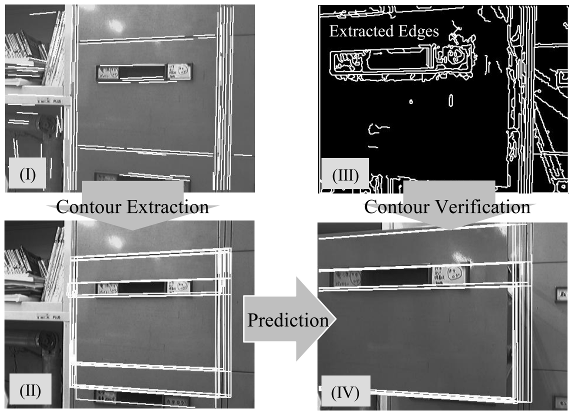
There are a lot of ref-LS in one image and each ref-LS has several candidates L_n. To select good correspondence in each ref-LS and to remove useless contour candidates, eq.(1) is used. That is, reprojection error is calculated from randomly selected L_n. If these L_n are well projected, good score is added to all of the selected L_n. This random selection and scoring are performed specified number, the best scored L_n is corresponded to each ref-LS. If the best score is worse than threshold, the ref-LS and its quadrangle is discarded. As a result, finally selected contours consist of proper contours.

D. Improvement of modeling

The more line segments exist, the more computation time is needed to correspondence procedure. So we propose two criteria for reducing front board candidates before calculating eq.(1).

Touched point of hand: Most of the furniture assumed in this paper have a knob in its front board area. Utilizing this rule, if front board candidates do not include a point which is the projection of the hand of manipulator, such candidates are discarded.

Appearance constraint: In our method, the quadrangle contours are extracted as front board candidates. Comparing with around environment, inner pixels of contours are invariant to the change through teaching if the contours indicate true front board. Utilizing this rule, making color histgram about each inner contours, only candidates which have similar state that of before and after teaching are adopted as true front board.



Before Teaching After Teaching

Fig. 6. Image Processing

V. Experiments

Fig.5, left shows our mobile manipulator system. In this experiment, cabinet was selected as target furniture. Fig.5, right shows direct teaching to the robot.

Fig.6 shows extracted contour candidates. Using our method, contour candidates can be selected from near to a front board. In this case, 81 candidates were extracted in initial process, but finally, 5 candidates were selected as a front board model. The processing time was 4.6 sec. We also tried to adopt our method to other types of furnitures (a sliding door and a small drawer) and achieved modeling as same as this experiment.

VI. Conclusion

We proposed a modeling method of furniture for a mobile robot. Without special tools or detail instruction, our method enables the robot to increase the handleable furniture. For modeling, direct teaching and generalized knowledge for handling are adopted. In this paper, we mentioned about modeling method of appearance and 3D shape information of furniture using external sensors. Experiment results show the effectiveness of our method.

References

[1] H. Asada and Y. Asari: "The Direct teaching of tool manipulation skills via the impedance identification of human motion", Proc. of IEEE Int. Conf. on Robotics and Automation, pp.1269–1274, 1988
[2] N. Y. Chong and K. Tanie: "Object Directive Manipulation Through RFID", Proc. Int. Conf. on Control, Automation, and Systems, Gyeongju, Korea, Oct. 22-25, 2003.
[3] R. Katsuki, J. Ohta, T. Mizuta, T. Kito, T. Arai, T. Ueyama and T. Nishiyama: "Design of Artificial Marks to Determine 3D Pose By Monocular Vision", Proc. 2003 IEEE Int. Conf. Robotics and Automation, pp.995–1000, 2003.
[4] J. Miura, Y. Shirai and N. Shimada: "Development of a Personal Service Robot with User-Friendly Interfaces", 4th Int. Conf. on Field and Service Robotics, pp.293– 298, 2003.
[5] L.Petersson, P. Jensfelt, D.Tell, M.Strandberg, D. Kragic and H.I. Christensen: "Systems Integration for Real-World Manipulation Tasks", Proc. 2002 IEEE Int. Conf. Robotics and Automation, pp.2500–2505, 2002.
[6] N.Sien, O.Stasse, Y.Kawai, T.Sakaguchi and K.Yokoi: "A Unified On-line Operation Interface for Humanoid Robots in a Partially-Unknown Environment", Proc. of IEEE Int. Conf. on Intelligent Robots and Systems, pp.4437–4439, 2006.
[7] http://www.hokuyo-aut.jp/products/urg/urg.htm