3D Shape Modeling of Movable Parts of Furniture Based on Time-Series Surface Correspondence

Kimitoshi YAMAZAKI [*1], Kiyohiro Sogen [*2], Takashi Yamamoto [*2], Masayuki Inaba [*3],

***Abstract*— This paper describes a modeling method that enables generation of a 3D shape model for movable parts of furniture and home appliances. Furniture such as refrigerators, shelves, and cabinets were targeted in this research. The proposed method is based on the tracking of 3D surfaces extracted from depth images captured by a 3D range camera. Both exterior and interior sides of an articulated part can be modeled as a 3D mesh model; moreover, the position of an articulated link can be estimated.**

I. INTRODUCTION

Robots working in everyday environments should have various types of abilities with respect to recognition, planning and manipulation, because people perform many possible daily tasks that can be performed by autonomous machines. This paper describes an automated object modeling method, that targets furniture and home appliance with articulated structures.

A simple and conventional way to provide object knowledge to robots working in everyday environments is to use approximated geometrical representation. In the past several decades, research studies have applied manually constructed geometric models that integrate surface color, texture, and articulated link [4], [6], [7], [9], [11], [13], [17]. This approach enables implementation of robust recognition and manipulation processes, however, manual model definition remains a problem. We have developed an assistive robot, the human support robot [3], to assist people in everyday activities and improve quality of life. We have also manually created a 3D geometric model of target objects and have used it to facilitate object detection and pose estimation; however, manual modeling was a burdensome task.

The purpose of this study was to construct a modeling method that targets furniture and home appliances, which often have articulated structures. For example, we targeted a refrigerator, as shown in Fig.1, because refrigerators have doors that must be opened. A robot must have a model of both sides of the door because the front side is required to detect the refrigerator from the surrounding environment. After opening the door, a model of the other side is required to search the storage space for an object. In this study, we aim to achieve automatic geometric modeling. The primary aim of this modeling method is 3D reconstructions of entire circumference shape of movable parts along an articulated link.

[*1] Faculty of Engineering, Shinshu University, 4-17-1 Wakasato, Nagano, Nagano, 380-8553, Japan. [*2] Toyota Motor Corp., Hirose, Toyota, Aichi, Japan. [*3] The University. of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan. kyamazaki@shinshu-u.ac.jp

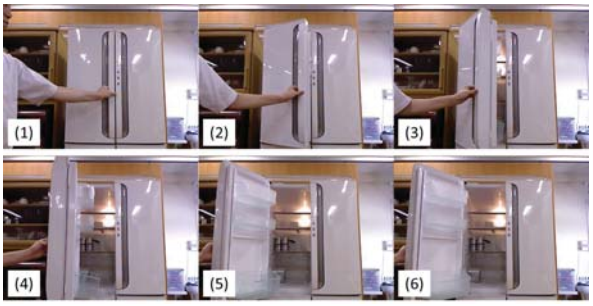

Fig. 1. Refrigerator door handled by an operator.

The remainder of this paper is organized as follows: Section II introduces related work and our approach. Sections III and IV explain the proposed method, and surface tracking and shape modeling, respectively. Section V presents experimental results, and conclusions are presented in Section VI.

II. RELATED WORK AND OUR APPROACH

A. Movable part model for a mobile manipulator

Many types of "mobile manipulator" robots have been presented, and some of them have attempted to handle movable parts with articulated structure existing in home environment. For example, Srinivasa et al. [13] developed a mobile manipulator that performs useful manipulation in a home. Their focus included opening a shelf door, which was casing grasp constraints. Meeussen et al. [6] performed door opening and closing using the PR2 robot. They integrated various technological elements such as grasp point recognition, arm planning, and footprint planning. Owing to the benefits of a soft manipulator, the PR2 achieved robust door manipulation. The abovementioned studies and many other studies provide a model of movable parts for a robot, however, the modeling burden was significant.

Several modeling methods have been proposed for objects with movable parts. Sturm et al. [14] performed modeling of an articulated door without markers on the door. This method is based on front board detection and tracking as well as a model of an articulated link. Yamazaki et al. [16] proposed the instructed motion model intended for drawers and doors with rotational, or translation axes. Observation data were obtained while an operator manipulated a robot directly, to model appearance, 3D shape, and a manipulation method. The abovementioned studies targeted only front board of a movable door. However, for some furniture and home

appliances, the inner side of the door is also important; e.g., the inner side of a refrigerator door shelving for food storage.

B. Our approach

We assume that input data for the proposed method is a sequence of depth images captured by a 3D range camera. The camera captures a situation in which an operator manipulates a movable part of a target object. From the data, 3D model of the entire circumference shape and an articulated part are reconstructed. Thus, to achieve effective modeling, we only need to show how to manipulate the target object. The procedure is simple compared to conventional approaches, in which shapes and articulated parts are modeled manually. This is the most significant advantage of the proposed method; however, there are two major difficulties which can be described as follows:

1) **Distinctive features cannot be extracted.**
   Our target objects, i.e., furniture and home appliances, often have textureless and shiny surfaces. If we aim to reconstruct 3D geometrical shapes using time-series sensor data, correspondence between two different sets of sensor data is needed. If we target objects with rich textures or characteristic shapes, we can use distinctive feature description [5] [12]; however, this is difficult for our case.
2) **A model should be unified both of surfaces.**
   As mentioned above, we aim to create a model that includes the entire circumference shape; thus, determining correspondences between sensor data is important. However, the appearance of a movable part in sensor data can change drastically while it is manipulated (See Fig. 1). One issue is determining how to find correct registration of sensor data captured for various views of a target object.

To overcome the abovementioned difficulties, we propose a surface-based modeling method. 3D surfaces are extracted from each sensor data, and their correspondences are investigated between two consecutive sensor data. Subsequently, surfaces belonging to a movable part are extracted, and a unified 3D shape is created; moreover, the position of an articulated link is also estimated.

III. Surface extraction and correspondence

A. Surface extraction

One approach to extract structural information from a 3D point cloud is plane detection. One study has proposed that modeling methods for daily environments must first assume that objects are formed by planar segments [15]. However, many daily objects have curved surfaces. To cope with such curved objects, we apply a region growing algorithm to detect locally independent regions.

In the algorithm, a starting point $\mathbf{p}_0$ is first selected. From this point, neighbor points whose similarity is over than predefined threshold are selected as homologous points. Our similarity measure is calculated by means of angle difference from neighbor normal vectors.

Fig. 2. Results of surface and contour detection

For this process, normal vectors should be calculated in advance for all points. Therefore, the following procedure is applied with an input image whose pixels have depth value d. Let a pixel p be a pixel of focus. A normal vector with respect to p is calculated from the 3D position of p and its neighbors. After performing calculations for the positional average and covariance matrix from the points, the direction and length of three orthogonal axes are acquired by means of eigenvalue decomposition. A normal vector is the axis whose length is the shortest at the three axes. Thus, one pixel has four variables, depth d, and normal vector $\mathbf{n} = (n_x, n_y, n_z)$.

For region growing algorithm, pixels are connected if the following rule is satisfied:

$$\begin{array}{ll} |d(i,j) - d(i+n, j+m)| & < d_{threshold}, \\ cos^{-1}(\mathbf{n}(i,j) \cdot \mathbf{n}(i+n, j+m)) & > \theta_{threshold}, \end{array} \quad (1)$$

where (i, j) denotes a coordinate on the present depth image, and $(i + n, j + m)$ are the coordinates of its neighbors. $(\cdot)$ indicates the calculation of inner product.

Region growing makes it possible to cope with curved surfaces, which differs from previous study that applied plane detection to represent a model [10]. Fig. 2 shows an example of detected surfaces that were extracted from an opened refrigerator. Each colored region indicates one surface, and white pixels within a region's boundary indicate a part belonging to a contour.

B. Surface correspondence

Many of surfaces are extracted from one depth image using region growing algorithm. To reconstruct the 3D shape, the correspondence of surfaces extracted from two consecutive images is required. To achieve this, criteria to evaluate the similarity of surfaces are also required.

Shape context [1] is utilized to generate a feature description for each surface. This provides low dimensional representation, a simple calculation process, and robustness to noise, which are suitable characteristics for our surface correspondence. The calculation procedure is as follows: First, 3D points that construct one surface are used for plane approximation by means of ellipsoidal approximation. Subsequently, 3D points belonging to a contour are projected onto the plane.

From these projected points, a frequency histogram is created by using the shape context description. In other words, voting plane is first prepared by radially dividing cells from the center of a surface. Subsequently, the number

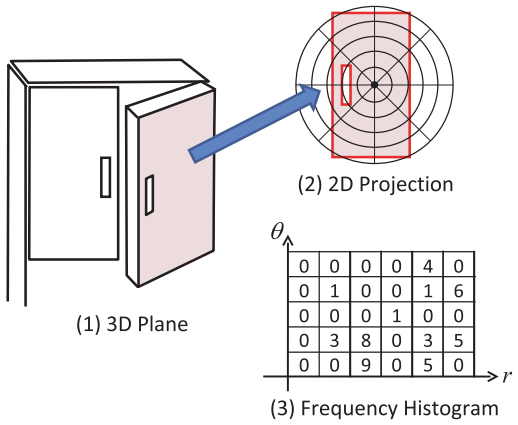

Fig. 3. Shape context

of contour points is counted at each cell (bin). The result is aligned in a two-dimensional histogram whose axes are a distance r from the center and angle θ of cell borders. In our case, r_{max}, which is upper limit of r, depends on the size of a focusing surface. Fig.3 shows this procedure.

Let $s_t^i \in \mathbf{S}_t = \{s_t^1, s_t^2 ... s_t^n\}$ be a list of surfaces extracted from sensor data at time t. Surface correspondence is performed at $t+1$ to find a surface s_{t+1}^j similar to s_t^i from another surface list $\mathbf{S}_{t+1} = \{s_{t+1}^1, s_{t+1}^2 ... s_{t+1}^m\}$. This is achieved by calculating the cosine similarity of two histograms. In the comparison process, the following criteria are used to filter out inadequate matching candidates in advance: (i) the Euclidean distance between s_t^i and s_{t+1}^j is not large, (ii) the size of r_{max} of s_t^i and s_{t+1}^j do not differ significantly. These criteria contribute to increased robustness and processing speed.

Because feature description by means of shape context is not rotation invariant, the starting θ position must be determined during histogram comparison. However, the description method has a structural difficulty; it produces discretization error when generating voting space. In addition, fluctuating sensor data is also a problem because measurement error is significant when using a Microsoft Kinect sensor. Moreover, differences in the pose of the observed object reduce repeatability of sensor data. To overcome this, we apply a marginalization process to a target cell and its neighbors in $r-\theta$ space, with the exception of a voting space whose value is determined according to the distance from the present r and θ.

IV. Entire Circumference Shape Modeling of Movable Part

A. Target surface selection belonging to movable part

Owing to the surface correspondence process, time-series relationships are represented as a set of sequence lists, which line up relationships for each surface between two consecutive measurement times. This result should be translated into another matrix representation that shows each surface has been observed from time $t=0$ to $t=T$. We refer to this matrix as the "surface tracking matrix". A list to which surface information is registered is required. For one surface, this list includes center coordinates, the point cloud, and the results of ellipsoid approximation.

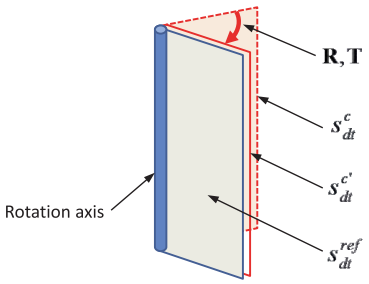

Fig. 4. Surface correspondence transformation

The next crucial issue for shape modeling is the extraction of surfaces belonging to a movable part. The procedure is described as follows. One surface is selected as reference surface. Let $\mathbf{S}^{ref} = \{s_{st}^{ref}, s_{st+1}^{ref}, ... s_{et}^{ref}\}$ be a reference surface that is selected from surface tracking matrix. st and et are the first and last time during s^{ref} existed. $\mathbf{S}^{ref}$ can be selected by determining the longest observed moving surface in the surface tracking matrix. Subsequently, the surface is evaluated by investigating its observed period, size, and motion according to the following equation:

$$\begin{aligned} &et - st > C_{time}, \\ &F_{area}(\mathbf{S}^{ref}) > C_{area}, \\ &F_{motion}(\mathbf{S}^{ref}) > C_{motion}, \end{aligned} \tag{2}$$

where C_{time}, C_{area}, and C_{motion} are predefined thresholds. F_{area} is a function that calculates the area of the approximated ellipsoid from the 3D points belonging to the target surface. F_{motion} is a function that calculates position transition between detection time and the end of tracking time. Each function processes time-series data as follows:

$$\begin{aligned} F_{area} &= \frac{1}{et-st}\sum_{t=st}^{et} A_{s_t}, \\ F_{motion} &= \max_{i,j\in[st,et]} |\mathbf{x}_i - \mathbf{x}_j|, \end{aligned} \tag{3}$$

where A_{s_t} denotes the area of surface s_t, and $\mathbf{x}_i$ denotes the center position of a surface at time i. The former equation ignores small surfaces; however, it does consider surfaces that are temporarily small with a certain level of penalty.

B. Adjunctive surfaces selection

The set of reference surfaces $\mathbf{S}^{ref}$ is used to evaluate whether other surfaces have moved with $\mathbf{S}^{ref}$. Let $\mathbf{S}^c$ be a set of surfaces to be evaluated, and td be the time when both $\mathbf{S}^{ref}$ and $\mathbf{S}^c$ existed.

A rotation matrix $\mathbf{R}$ and a translation vector $\mathbf{T}$, which show a coordinate transformation between s_t^{ref} and s_{td}^{ref} ($t < td$), are calculated. A rotation matrix is defined for each surface as the normalized three major axes of an approximated ellipsoid; thus, $\mathbf{R}$ is calculated as the relative

rotation matrix between rotation matrices at time t and td. The translation vector is the displacement between $\mathbf{x}_t^{ref}$ and $\mathbf{x}_{td}^{ref}$ and is the center coordinates of the ellipsoid. $\mathbf{R}$ and $\mathbf{T}$ are used to transform s_t^c to $s_{td}^{c'}$. If $s_{td}^{c'}$ is well superimposed to s_{td}^c, $\mathbf{s}^c$ and $\mathbf{s}^{ref}$ are considered in the same group because to a single body.

Measurement error introduces harmful effect; thus, judging whether a set of surfaces $\mathbf{S}^c$ and a set of surfaces $\mathbf{S}^{ref}$ belong to the same group may be difficult if we use only a single s_t^c value. Therefore, we employ multiple sampling during time st to et. Sampling results are used to evaluate a surface group $\mathbf{s}_t^c$, and the results are summed up. The evaluation divides surfaces into three types: (i) movable part, (ii) background, and (iii) others. An example of item (iii) is a human arm that is manipulating the target furniture.

The basic equation is as follows:

$$\{movable, background, others\}^c = f(s_t^c, s_{td}^c, s_{td}^{c'}), \quad (4)$$

where $f(\cdot)$ is a function to obtain a maximum value from the evaluation items C_m, C_b, and C_o.

$$\begin{array}{l} C_m = \sum_t g_m(s_t^c, s_{td}^c, s_{td}^{c'}), \\ C_b = \sum_t g_b(s_t^c, s_{td}^c, s_{td}^{c'}), \\ C_o = \sum_t g_o(s_t^c, s_{td}^c, s_{td}^{c'}). \end{array} \quad (5)$$

The following equation is used for movable part,

$$g_m(s_t^c, s_{td}^c, s_{td}^{c'}) = w_{dist}^m D_{c'-td/c-td}^{dist} + w_{ang}^m D_{c'-td/c-td}^{ang}, \quad (6)$$

where w_{dist}^m and w_{ang}^m are weight coefficients that are tailored to movable parts. For function g_b for a background component, $c' - td/c - td$ in the above equation is changed to $c' - td/c - t$. g_o is a function to receive a value from a function D when both g_m and g_b become 0, which is expressed as follows:

$$D_{a-p/b-q}^* = \begin{cases} 1 & if \;\; h^*(s_p^a, s_q^b) < threshold \\ 0 & otherwise, \end{cases} \quad (7)$$

where $*$ is $dist$ or ang. $h^{dist}(x, y)$ is a position similarity function based on Euclidean distance, and $h^{ang}(x, y)$ is a normal vector similarity function based on the inverse of cosine similarity.

C. The creation of entire circumference shape model

Surfaces considered as movable parts can be extracted as a time-series data set using the methods discussed in Sections IV-A and IV-B. Using these data, a shape model that concatenates the surfaces into a single body is generated according to the following procedure.

First, canonical time tb is defined, and a surface s_{tb}^{ref} to be focused on is determined. Several sensor data at time t are selected as key frames; they capture a movable part from the front, back, and lateral sides. Pose transformation between s_t^{ref} and s_{tb}^{ref} is calculated as $\mathbf{R}_{t-tb}$ and $\mathbf{T}_{t-tb}$. Subsequently, a group of surfaces $\mathbf{S}_t^c$ considered as portions of a movable part are transformed by $\mathbf{R}_{t-tb}$ and $\mathbf{T}_{t-tb}$. This manipulation produces a group of surfaces $\mathbf{S}^m$ that exist simultaneously (This concept is the same as $s_{td}^{c'}$, which was described in former section).

However, $\mathbf{S}^m$ may lack some surfaces belonging to the movable part. For instance, the extraction of movable surfaces becomes unstable against a markedly uneven region; thus, sufficient shape data may not be obtained after successful 3D measurement. To overcome this problem, a rectangular box that includes all movable surfaces $\mathbf{S}^m$ is created. If there are other surfaces in the rectangular box, they are also considered a movable part.

The rectangular box is created by means of principle component analysis (PCA). However, PCA using 3D point coordinates may produce an inclined box from a desired pose if the positions of the target surfaces are biased. To overcome this, normal principle component analysis (NPCA) [8] is applied. Normal components embedded in the point cloud of target surfaces are used to obtain the main axis from the following covariance matrix:

$$\mathbf{C} = \sum_{i=0}^{N} \mathbf{n}_i \mathbf{n}_i^t, \quad (8)$$

where N is the number of normal components used for NPCA. Here $\mathbf{n}$ denotes a 3×1 normal vector, which is calculated by the procedure explained in Section III-A. Each surface of a rectangular box is defined as perpendicular to the axis calculated by NPCA, and the size of the box is determined from the maximum and minimum 3D point coordinates belonging to the target surfaces.

Original time-series depth images are used to search inlier 3D points of the box. Subsequently, triangular meshes are created by connecting neighboring points. Finally, a dense 3D shape model is obtained.

D. Estimation of articulated part: pivoted door

The time-series $\mathbf{S}^{ref}$ is used to estimate an articulated link. First, a list of coordinates $\mathbf{X} = \{\mathbf{x}_{st}, \mathbf{x}_{st+1}, ...\mathbf{x}_{et}\}$ that describes the center of surfaces $\mathbf{S}^{ref}$ from time st to et is collected. Subsequently, difference vectors $\mathbf{V} = \{\mathbf{v}_{st}, \mathbf{v}_{st+1}, ...\mathbf{v}_{et-1}\}$, where $\mathbf{v}_i = \mathbf{x}_{i+1} - \mathbf{x}_i$, are calculated. Then, planes that are perpendicular to $\mathbf{v}_i$ and pass through $\mathbf{a}_i = \mathbf{x}_{st+i} + 0.5 * \mathbf{v}_{st+i}$ are calculated. Finally, lines $\mathbf{L} = \{l_{st}, l_{st+1}, ...l_{et-1}\}$ are obtained by calculating the intersection of two planes extracted from the above planes.

A line l is selected, and then the distance from $\mathbf{X}$ is calculated. If these distances are approximately equal, the line is considered an approximation of the center axis of a cylinder, i.e., rotational axis of the door.

V. Modeling Experiments

A. Experimental settings

A Microsoft Kinect 3D range camera was used to assess the effectiveness of the proposed method. The size of the depth image was 640×480 pixels.

A refrigerator and a drawer were selected as target objects. The range camera was fixed in front of these target objects, and an operator opened their movable parts. For the drawer case, the range camera was moved around it after the opening operation. Then, a model was generated from a set of depth images.

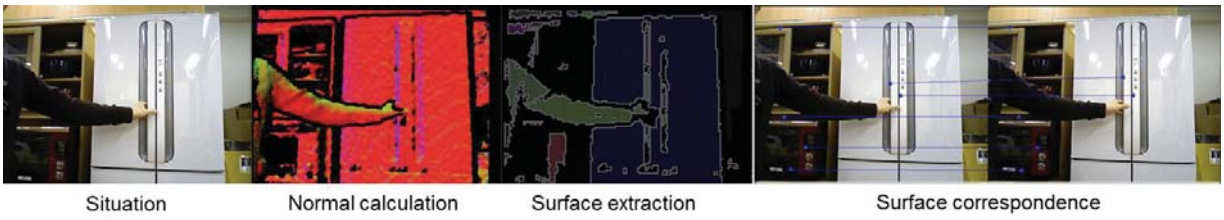

Fig. 5. Contour detection and tracking

B. Region growing

The region growing algorithm described in section III-A was applied to an input image. First, normal information was calculated. A depth value $d(i, j)$ and its two neighbors $d(i+n, j)$ and $d(i, j+m)$ were used to create a triangle patch. The normal vector of the triangle patch was assigned to the (i, j) pixel. In this experiment, we set $n = m = 4$. After each normal calculation, smoothing process was added using neighbor normal vectors near $n \leq 4$ and $m \leq 4$. The "normal calculation" shown in Fig. 5 represents a single result. The region growing algorithm was then applied using this result. The "surface extraction" in Fig. 5 shows an example. Region growing enabled us to obtain fairly correct regions despite having to cope with curved surfaces such as the refrigerator door. It is difficult to use plane detection method.

We used a computer with a 2.67 GHz, 4 core CPU, and the abovementioned process required about 100 to 110[ms/frame] to generate surfaces.

C. Surface correspondence

The pair of images shown in Fig. 5 is the result of surface correspondence. The blue point shows the center of each surface, and the blue line shows the surface correspondence between two consecutive frames. This process, including descriptor calculation, required 20 – 40 [ms] depending on the number of surfaces that required correspondence calculation.

We used 12 image streams that captured the refrigerator from different viewpoints, and the average number of correspondences per pair was 24.3. For the surface correspondence, although there were nearly no false positives, the correspondence sequence was sometimes broken when the region growing algorithm originally divided one surface into two surfaces. This tended to occur when the angle between the optical axis and a surface became large. An angle of more than 60 [deg] causes a lack of measurement data with the case of Kinect sensor.

D. Modeling results: refrigerator door

Fig. 6 and Fig. 7 show the results of the refrigerator door modeling. To make surface correspondence, 98 frames were used, and surfaces belonging to movable part were selected according to the methods discussed in section IV-A. Examples of refrigerator surfaces are shown in Fig. 6, in which 17 surfaces were regarded as movable parts.

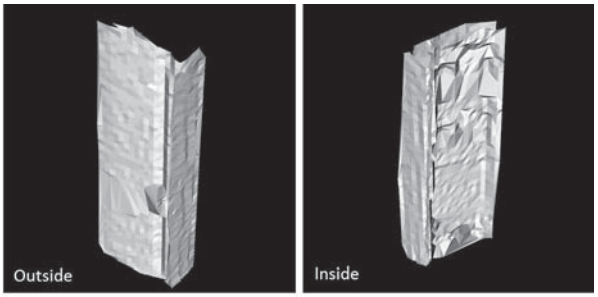

Fig. 6. 3D model of the refrigerator door

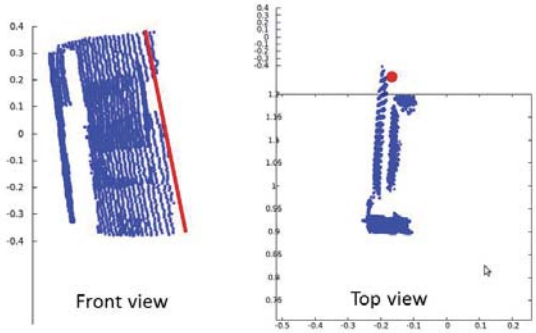

Fig. 7. Estimated rotation axis

Depth images selected with 10 frames were considered key frames, and the 3D points composing these surfaces were used to generate a rectangular box. Then, other 3D points included in the box were added. In each key frame, 2D Delaunay triangulation was performed against pixels belonging to the surfaces of movable parts, and the 3D coordinates of vertices were reconstructed. Fig. 6 shows an example. Note that the middle part of the front side of the door could not be measured because it was occluded by an operator arm. However, from the 3D mesh interpolation, we obtained a surface model. This modeling process took approximately 1.0 s.

Fig. 7 shows rotation axis of the door. Red line shows the estimated position of the axis.

E. Modeling results: drawer

First, the surface-based modeling described in sections III, IV-A, and IV-B was performed using depth images that

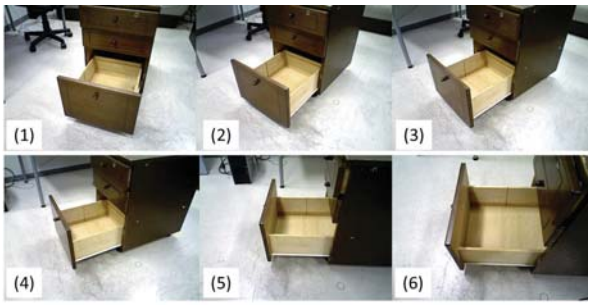

Fig. 8. Multi-view measurement after opening a drawer.

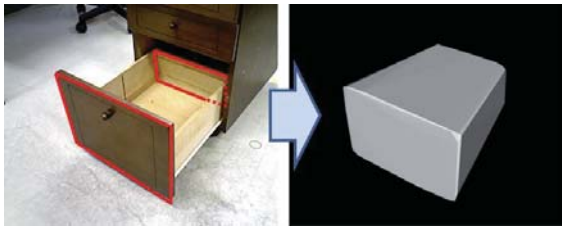

Fig. 9. Convex hull generated from drawer surfaces

captured a drawer opened by an operator. In this phase, a 3D range camera was fixed. The 3D points belonging to surfaces extracted as movable parts were used to generate a convex hull (Fig. 9). The red rectangles in Fig. 9 indicate surfaces selected as movable parts. There were used to generate the convex hull shown in the Fig. 9. The 3D points in the convex hull were considered as belonging to the movable part.

To cope with drawer type furniture, we applied additional multi-view observation. A point-to-point registration method was performed using these depth images. Iterative Closest Point (ICP) algorithm [2] was applied. was applied to 56 depth images to produce the final model, which is shown in Fig. 10.

VI. Conclusions

In this paper, we have proposed a method of 3D shape modeling. Our main targets were furniture and home appliances which often have articulated structure. We aimed to achieve automatic geometric modeling whose primary role is 3D reconstruction of entire circumference shape of movable part along an articulated link. We have proposed a surface-based modeling method in which 3D surfaces are extracted from sets of sensor data, and their correspondences are investigated between two consecutive sets of data. Then, surfaces belonging to a movable part are extracted and a unified 3D shape is created. The position of an articulated link was also estimated.

In future, we plan to implement the proposed method as an application in a previously proposed home assistant robot [3].

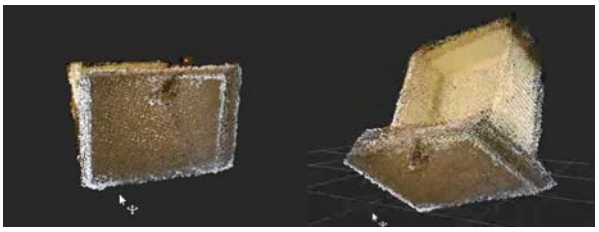

Fig. 10. Drawer shpae model

References

[1] S. Belongie, J. Malik, and J. Puzicha: “Shape Matching and Object Recognition Using Shape Contexts,” IEEE Transactions on Pattern Analysis and Machine Intelligence vol. 24 No. 24, pp. 509–521, 2002.

[2] Besl, Paul J., N.D. McKay: “A Method for Registration of 3-D Shapes,” IEEE Trans. on Pattern Analysis and Machine Intelligence Vol. 14, No.2, pp. 239–256, 1992.

[3] K. Hashimoto, F. Saito, T. Yamamoto and K.Ikeda: “A Field Study of the Human Support Robot in the Home Environment,” in Proc. of IEEE Workshop on Advanced Robotics and its Social Inpacts, pp. 143 – 150, 2013.

[4] K. Nagatani and S. Yuta: ”Designing strategy and implementation of mobile manipulator control system for opening door,” IEEE Int’l Conf. on Robotics and Automation, Vol 3, pp 2828 – 2834, 1996.

[5] D. G. Lowe: “Distinctive image features from scale-invariant keypoints,” Int’l J. of Computer Vision, vol. 60, No. 2, pp. 91-110, 2004.

[6] W. Meeussen, et al.: “Autonomous door opening and plugging in with a personal robot,” in Proc. of IEEE International Conference on Robotics and Automation, pp. 729 – 736, 2010.

[7] K. Okada, M. Kojima, S. Tokutsu, T. Maki, Y. Mori and M. Inaba: “Multi-cue 3D Object Recognition in Knowledge-based Vision-guided Humanoid Robot System,” Proc. of the IEEE/RSJ Int’l. Conf. on Intelligent Robots and Systems, pp.1505–1506, 2007.

[8] P. Papadakis, I. Pratikakis, S. Perantonis, T. Theoharis “Efficient 3D shape matching and retrieval using a concrete radicalized spherical projection representation,” Pattern Recognition, vol. 40, no. 9, pp. 2437 – 2452, 2007.

[9] L. Petersson, P. Jensfelt, D. Tell, M. Strandberg, D. Kragic, and H. I. Christensen: “Systems integration for real-world manipulation tasks”, in Proc. Int. Conf. on Robotics and Automation, vol. 3, pp. 2500 – 2505, 2002.

[10] J. Poppinga et al.: “Fast Plane Detection and Polygonalization in Noisy 3D Range Images,” Int ’ l Conf. on Intelligent Robots and Systems, 2008.

[11] U. Reiser, et al.:”Care-O-bot 3 – Creating a product vision for service robot applications by integrating design and technology,” in Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1992-1997, 2009.

[12] R. B. Rusu, Nico Blodow and Michael Beetz: “Fast Point Feature Histograms (FPFH) for 3D registration,“ IEEE International Conference on Robotics and Automation, pp. 3212 - 3217, 2009.

[13] S. Srinivasa, D. Ferguson, C. J. Helfrich, D. Berenson, A. Collet, R. Diankov, G. Gallagher, G. Hollinger, J. Kuffner, M. Vande Wegh: “HERB: A Home Exploring Robotic Butler,”, Journal of Autonomous Robots, Vol.28, pp.5 – 20, 2009.

[14] J. Sturm, A. Jain, C. Stachniss, C. Kemp and W. Burgard: “Robustly Operating Articulated Objects based on Experience,” in Proc. of Int ’ l Conf. on Intelligent Robots and Systems, 2010.

[15] J. Weingarten and R. Siegwart: “EKF-based 3D SLAM for Structured Environment Reconstruction,” in Proc. of Int’l Conf on Intelligent Robots and Systems, pp. 3834 - 3839, 2005.

[16] K. Yamazaki, T. Tsubouchi and M. Tomono: “Furniture Model Creation through direct teaching to a Mobile Robot,” Journal of Robotics and Mechatoronics, vol.20, no.2, pp.213-220, 2008.

[17] K. Yamazaki, R. Ueda, S. Nozawa, M. Kojima, K. Okada, K. Matsumoto, M. Ishikawa, I. Shimoyama and M. Inaba: “Home Assistant Robot for an Aging Society,” in Proc. of the IEEE, Vol.100, No.8, pp.2429 – 2441, 2012.