Articulated Object Modeling based on Visual and Haptic Observations

Wei Wang[1], Vasiliki Koropouli[1], Dongheui Lee[1] and Kolja Kühnlenz[2,3]
[1]*Institute of Automatic Control Engineering (LSR), Technische Universität München, D-80290 München, Germany*
[2]*Institute of Advanced Study (IAS), Technische Universität München, D-80290 München, Germany*
[3]*Bayerisches Landesamt für Maß und Gewicht, D-80638 München, Germany*

Keywords: Articulated Object Modeling, Object Skeletonization, Vision-based Articulated Object Manipulation.

Abstract: Manipulation of articulated objects constitutes an important and hard challenge for robots. This paper proposes an approach to model articulated objects by integrating visual and haptic information. Line-shaped skeletonization based on depth image data is realized to extract the skeleton of an object given different configurations. Using observations of the extracted object's skeleton topology, the kinematic joints of the object are characterized and localized. Haptic data in the form of task-space force required to manipulate the object, are collected by kinesthetic teaching and learned by Gaussian Mixture Regression in object joint state space. Following modeling, manipulation of the object is realized by first identifying the current object joint states from visual observations and second generalizing learned force to accomplish the new task.

1 INTRODUCTION

Most tasks in human daily life require manipulation of articulated objects of one or more degrees of freedom. Some characteristic examples of such tasks consist of door opening, drawer pulling and rotating a water tap. Manipulation of articulated objects is a great challenge for robots which are required to recognize an articulated object mostly by vision and make a decision about how to manipulate it. By making robots capable of manipulating articulated objects, they could enter more actively human life and help humans with dangerous or difficult tasks as well as helping elderly people in daily life.

Many previous works on articulated object modeling mainly focus on solving the problem of identifying the kinematic characteristics of articulated objects using different types of sensor systems. In (Sturm et al., 2011), an approach is presented to learn kinematic models of articulated objects from observations, which does not allow for object identification, and ignores kinematic joint localization and constrains in object. In (Katz and Brock, 2008), kinematic task-relevant knowledge is acquired and learned in object's joint state space. This is realized via interaction with the environment and, finally, a kinematic model of the object is incrementally built. However, only visual data is employed and information about the dynamic properties of the object is not taken into account for manipulation. In (Huang et al., 2012), joint axes' position of an articulated object is estimated given different object configurations from depth image data. This aims at providing the grasp-

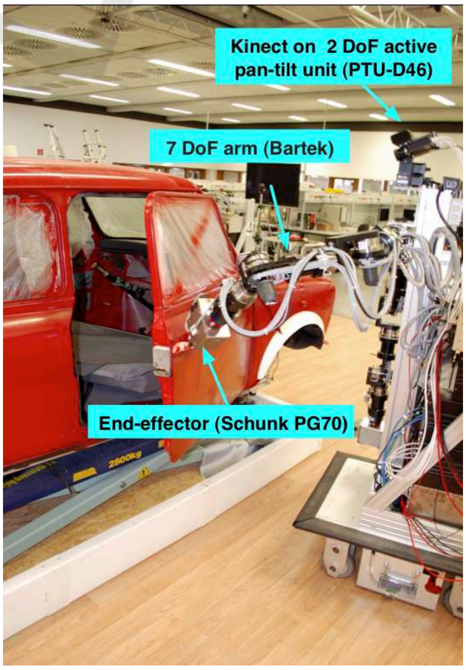

Kinect on 2 DoF active pan-tilt unit (PTU-D46)
7 DoF arm (Bartek)
End-effector (Schunk PG70)

Figure 1: A 7 DoF robotic arm manipulates a car's door (single revolute joint articulated object).

Wang W., Koropouli V., Lee D. and Kühnlenz K..
Articulated Object Modeling based on Visual and Haptic Observations.
DOI: 10.5220/0004280902530259
In *Proceedings of the International Conference on Computer Vision Theory and Applications* (VISAPP-2013), pages 253-259
ISBN: 978-989-8565-48-8

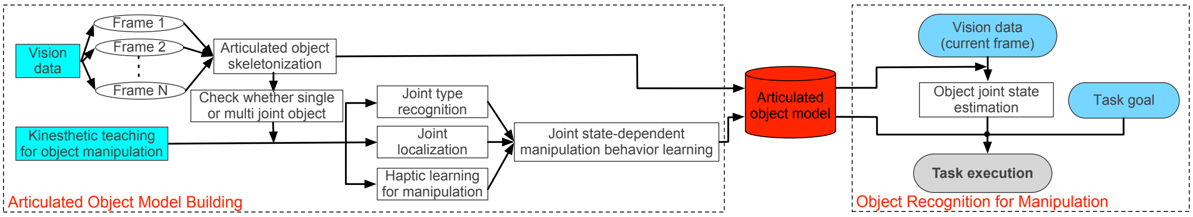

Vision data
Frame 1
Frame 2
Frame N
Articulated object skeletonization
Check whether single or multi joint object
Joint type recognition
Joint localization
Haptic learning for manipulation
Joint state-dependent manipulation behavior learning
Kinesthetic teaching for object manipulation
Articulated Object Model Building
Articulated object model
Vision data (current frame)
Object joint state estimation
Task goal
Task execution
Object Recognition for Manipulation

Figure 2: Proposed framework.

ing point and position trajectory to the robot. However, recognition of the object configuration is not considered. All previous works lack a framework for recognition of different articulated objects. In addition, they do not allow to estimate the current joint states of the object and adapt the manipulating behavior accordingly. In addition, previous works do not account for learning the force that is required to operate an object. For example, opening a completely closed or semi-closed door are two different tasks which require different manipulating forces. Some other works focused on learning manipulation of articulated mechanisms by learning force control skills while ignoring the mechanism structure of the object. In all these works (Kalakrishnan et al., 2011), (Lutscher et al., 2010), no visual information is used to recognize the object and characterize the number and type of joints and the constraints that apply on each joint of the object. Therefore, these approaches cannot generalize to the objects with different structures or configurations.

All works on articulated objects so far focus, either on using visual data for object characterization without learning manipulation force, or on learning manipulation force skills without analyzing the articulation characteristics of the object. Learning manipulation of even a single-joint articulated object is a challenging problem, since the articulation characteristics of the object have to be extracted first before appropriate manipulation force is learned. We thus, first seek to solve the problem for single-joint articulated objects and extend in future works to multiple-joint objects. In this paper, a framework for learning manipulation skills for single-joint articulated objects is proposed, which consists of (a) skeletonization of object, (b) joint number estimation based on object skeleton trace from different visual frames, (c) characterization of joint type, and (d) learning of Cartesian force which is required for manipulation. In particular, visual data are employed to build the object skeleton and estimate the current state of the object's joint. The trace of the skeleton nodes over time is employed to determine whether it is a single- or multi-joint object. In addition, haptic data in the form of Cartesian-space forces are captured from multiple human demonstrations by kinesthetic teaching and learned in object's joint state space. Generalization of manipulation force can be realized based on current joint's state and the task goal.

This paper is organized as follows. In Section 2 We define our problem and propose a method for skeletonize an articulated object and learning the manipulation force. In Section 3, the experimental setup and results are presented.

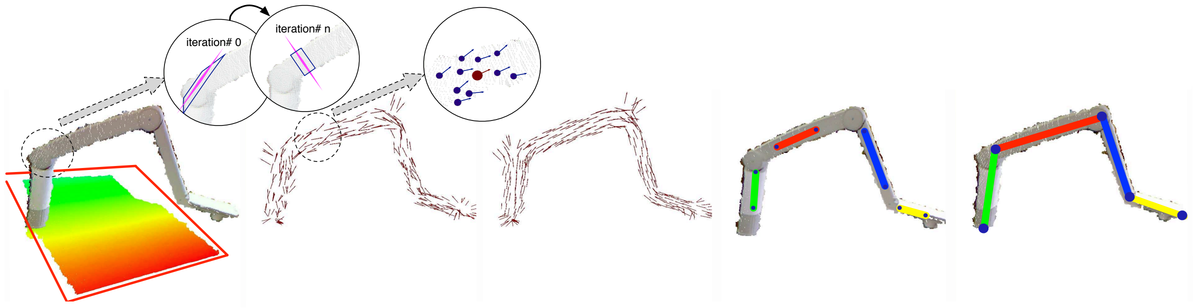
2 PROPOSED APPROACH

To manipulate articulated objects, information about both the structure and the kinematic and dynamic properties of the object is required. An articulated object could be described by its number and type of joints, link properties and kinematic relationships between neighboring links. Basic geometry features which are used for rigid object modeling and recognition, such as Viewpoint Feature Histogram (VFH) (Rusu et al., 2010), are not suitable for deformable objects. However, these approaches require complete depth information of the object. Since articulated objects can lie in a practically huge number of different configurations, capturing information about all these potential configurations is practically infeasible. For this reason, object skeletonization is the most suitable method for extracting the structure and kinematic constraints of an object. We define the model of an articulated object as

$$Obj = (S,\ J_m(T,\ P,\ C),\ \boldsymbol{f}), m = 1,..,M \qquad (1)$$

where S represents the skeleton of the object which is used for object recognition, J_m joint descriptor of the m-th joint, T joint type, P joint position and C joint constraints. The $\boldsymbol{f}(J_1,...,J_M)$ is the Cartesian force which is needed to manipulate the object where $J_1,...,J_M$ are joint descriptors of the articulated object where M is the number of joints.

Investigating multiple-joint objects is highly complicated and implies sufficient modeling of all individual joints of the object. For this, in this paper, we focus on modeling of single-joint articulated objects where visual and haptic information is integrated for



a) Object extraction based on RANSAC plane estimation

b) Vector field generation based on optimized cutting plane

c) Vector field smoothing based on Gaussian-weighted nearest neighbors

d) Line shape skeleton extraction based on vector field clustering

e) Skeleton topology estimation based on the line growth

Figure 3: Skeletonization steps of a multi-joint articulated object (phone arm).

highly efficient object manipulation. The framework presented here can be extended to modeling multiple-joint objects though and this is going to be presented in future work. Manipulation force constitutes part of an object's model since it indicates the dynamic properties of the object. This force is critical to the success of a robotic task and depends on the object's current joint states. The manipulating force can be represented by $\boldsymbol{f} = \pi(s_{J_m},\ e),\ m = 1,\ldots,M$, where π is a force generation policy, s_{J_m} the state of the m-th joint which may represent the angle of a rotational joint or length of a prismatic joint and e the task goal.

Fig. 2 shows the framework which is used to model a single-joint articulated object. The framework consists of two main components which are building a database of articulated objects' models and recognizing an incoming object based on visual and haptic information. The modeling stage can be divided into two parts where the first part involves vision-based object skeleton extraction and the second part consists of identification of the object's dynamic properties by teaching the robot appropriate force to operate the object. The kinematic joint properties (T,P,C) of a joint J are estimated from observation of the skeleton S across multiple configurations. Using learning by demonstrations, the appropriate force $\boldsymbol{f}$ is learned in the object's joint space. During generalization, the robot observes the object and extracts its current joint state. The force is generated based on the task goal such as the position or joint angle the object should finally reach and its current joint state.

2.1 Object Skeletonization

A point cloud, in terms of depth image data of an object, is used for skeletonization of the articulated objects. This is realized by observing multiple frames of the object's kinematic links. The skeleton of the object is extracted which allows to recognize the object and estimate its current joint states. Based on extracted object skeleton and the location of skeleton nodes, the object is classified as a single or multi-joint object. Skeleton models which represent the medial axis of a 3D model are widely used for object reconstruction and arterial object analysis. In (Tagliasacchi et al., 2009), *rotational symmetry axis* is used for the object skeleton points estimation. This work requires the full range point cloud of the object and uses the assumption that all object's model should be pipe-like. Instead, in this paper, a novel method of skeletonization of articulated objects is presented, which is not based on pipe-like configurations only but it can identify objects of abstract structures such as plane-like structure. The phone arm shown in Fig. 3 and car's door shown in Fig. 5 are two examples of objects with different type of structure, the first pipe-like and the latter plane-like.

2.1.1 Vector Field Generation

Firstly, the Random sample consensus (RANSAC)-based plane fitting algorithm is used to extract the object point cloud from the background (Rusu et al., 2010), shown in Fig. 3(a) and Fig. 5(b). The vector field presents the best local rotational symmetry of each point in the extracted object point cloud. Our method extracts the vector field using the optimized cutting plane. Based on RANSAC plane estimation with a certain number of iteration steps T_c, the vector field over the data points is generated. The best cutting plane $C_c = plane[x_i, v_i]$ which goes through the point x_i with the normal $\hat{v}$ is estimated by minimizing the number of inliers which are within the distance d_c. In addition, these points should also be in the same cluster N_i of the related point x_i using the geometric nearest neighbors:

$$\hat{v}_i = \underset{v\in\Re^3,\|v\|=1}{\arg\min}\ num(\{j_{N_i}|\ \|c_j - C_c^{(t)}\| \le d_c; x_j \in X_{raw}\}),$$

where $t \in [1, T_c]$ is the iteration index. Fig. 3(b) shows the result where the circles show the iteration step.

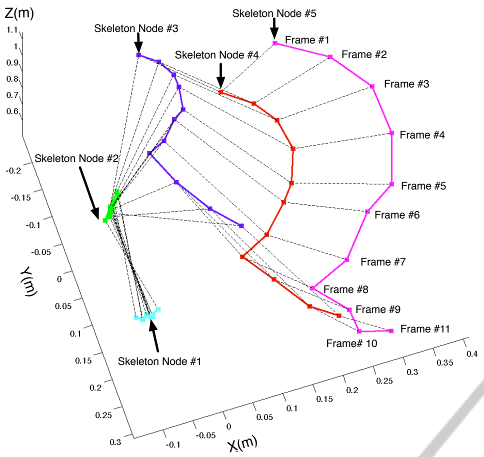


Figure 4: Skeleton node traces through different visual frames: black lines present the skeleton topology; each skeleton node trace is shown by a different-color solid line.

Note that, the direction of the optimized cutting plane could be the inverse which, however, will not influence the final results. The directions are reorganized based on the base plane coefficients.

A Gaussian-weighted method is developed for the vector field smoothing. The point x_i with normal v_i has the neighbor cluster X_i with points number n, which is determined by the distance threshold d_s. The weight function w is defined based on the gaussian contribution, decided by each neighbor's 3-D distance respect to the point x_i:

$$\begin{aligned} w_j &= \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{1}{2\sigma^2}\|x_j - x_i\|^2), \\ v_{i:new} &= \frac{\sum_{j=1}^{n} w_j v_j}{\sum_{j=1}^{n} w_j},\ x_j \in X_i. \end{aligned} \tag{2}$$

In our case, the standard deviation $\sigma = 1$ is used. Fig. 3(c) and Fig. 5(c) shows the smoothed vector fields over the object in different shapes.

2.1.2 Line-shape Skeleton Estimation

The skeleton of the object is described with the lines and linked nodes named skeleton nodes. After smoothing, the vector field is clustered using the nearest neighbor clustering method (Wang et al., 2011), which considers the positions and the directions. Meanwhile, the final skeletal point position could be extracted using the centering of the raw object points, which should be in the cutting plane through the related vector point with distance threshold. These skeletal points could be extracted from planar object. Instead, (Tagliasacchi et al., 2009) minimize the sum of squared distances from the point to the related normals, which will cause the position of the skeletal points for the planar object become infinite. The best line l could be extracted to minimize the distance sum from the extracted skeletal points. The line detection result is shown in the Fig. 3(d).

2.1.3 Skeleton Topology Extraction

The line detection result presented in Fig. 3(d), does not constitute the whole skeleton of the object since some skeleton points have been filtered out by clustering step. For this, the line growth algorithm is used to estimate the whole skeleton topology. All the detected lines grow in both positive and negative direction to overcome the whole skeleton. The lines stop growing when they,

(i) reach the edge of the object point cloud and are viewed as skeleton root node as the Node 1 and Node 5 in Fig. 4;

(ii) meet another skeleton line and at that time they stop growing up and are characterized as skeleton link node as the Node 2, 3 and 4 in Fig. 4.

These points are clustered and merged using 3-D Euclidean clustering (Wang et al., 2011). Then whole object skeleton nodes are extracted. Meanwhile the root and link nodes indicate the topology of object skeleton. The results are shown in Fig. 3(e) and Fig. 5(c). Different colored points represent the different estimated skeleton nodes and the dashed line links represents the skeleton topology.

2.1.4 Kinematic Joint Number Determination

As shown in Fig. 4 and Fig. 5(d), the object skeleton topology is extracted frame by frame with different configurations of articulated objects. The dashed lines represents the object skeleton topology and the traces of different extracted skeleton nodes are shown as different colored solid line. With the traces of skeleton nodes with different frames, all the dynamic observations are obvious. From frames 1 to 8, it is obvious that the observation patterns of nodes 3 to 5 differ from the patterns from frame 8 to 11. These two kinds of patterns in terms of the skeleton topology of object are changing, imply that the estimated object is not the single joint articulated object. The skeleton node S_i with index i is viewed as the base node to estimate the Euclidean distances with others as $E_i = \|S_0 - S_i\|, i \in [1, n]$, which is used to calculate the difference cost function DIF_j between current frame j with the previous frame $j-1$ as following:

$$DIF_j = \sum_{i=1}^{n} \frac{|E_i^j - E_i^{j-1}|}{E_i^{j-1}},\ j \in [1,\ F] \tag{3}$$

where F is the number of frames. At the frame 9, DIF_9 increased significantly, which means this articulated object contains multi kinematic joints. In comparison, as shown in Fig. 5(d), the door of car is the single joint articulated objects.

With the certification of the joint number from the object skeleton topology observations of different demonstrations, the kinematic joint characterization and localization could be extracted in the different strategies. As the one joint articulated object, the trajectory from one of object skeleton nodes could represent the whole object motion pattern and be used for its kinematic joint characterization. Otherwise, for multi joint articulated object, we need to analyze all the skeleton nodes trajectories hierarchically to extract all the kinematic joints' properties.

2.2 Kinematic Joint Characterization

The kinematic joints of the articulated object are distinguished into two types, prismatic and revolute (Sturm et al., 2011). Given the positional trajectories of the end-effector of the object, it is rather straightforward to discriminate between the two types of joints. The position vector of the point A of an articulated object which is moving in the 3D space can be expressed by $\vec{g} = g_x\hat{x} + g_y\hat{y} + g_z\hat{z}$. If only one positional component is non-zero, the joint is prismatic. The positional components are digitized as follows: if a component is different than zero, it is assigned the value 1, else the value 0. The digitized components g_x, g_y and g_z can the input to a Boolean logic scheme which is equivalent to the numerical computation given by

$$Y = (g_x + g_y + g_z - g_x g_y g_z)\,(g_x + g_y - g_x g_y)\,. \quad (4)$$

By applying (4) at each time step and taking the average $\overline{Y}$ of all outputs $Y(n)$ where n is the time index, we deduce whether the joint is revolute or prismatic. If $\overline{Y} = 0$ then the joint is prismatic. If $\overline{Y} \neq 0$, the joint is revolute. In case that a joint is revolute, and thus, causes a rotational movement, the angle range of the joint is estimated. The positional data of the end-effector of an articulated object are recorded during demonstrations of the task. The angle range is computed by $\theta(n) = \arctan(\overline{g}_i(n)/\overline{g}_j(n))$, where $n = 1, \ldots, N$ is the time index and $\overline{g}_i$ and $\overline{g}_j$ the two non-zero average positional trajectories in directions i and j. The average positional trajectories are computed, since many demonstrations are available, as $\overline{g}_i(n) = \frac{1}{K}\sum_1^K g_i^{(k)}$, $\overline{g}_j(n) = \frac{1}{k}\sum_1^K g_j^{(k)}$, where $g_a^{(b)}$ is the position of demonstration b in direction a and K is the number of demonstrations of the task.

2.3 Learning Force Skills

We desire to extract an average expert behavior for a task based on multiple demonstrations (Lee and Ott, 2011). Since the speed of the demonstrator varies from trial to trial and demonstrations are not time-aligned, demonstrations become time-aligned by Dynamic Time Warping. The force policy of a task is extracted from multiple demonstrations using a probabilistic approach proposed in (Calinon et al., 2007). This approach consists of Gaussian Mixture Modeling and Regression and estimates a smooth generalized version of demonstrated signals which captures all the important features of the task.

Time-aligned data pairs $d_i = \{s_i, \boldsymbol{f}_i\},\ i = 1, \ldots, N$ are considered, where N is the number of data points in each demonstration, s_i the input joint states and $\boldsymbol{f}_i \in \Re^{D\times N}$ represent force data where D is the dimensionality of $\boldsymbol{f}$. A mixture of L Gaussian functions is considered with probability density function $p(d_i) = \sum_{l=1}^{L} p(l)p(d_i|l)$, where $p(d_i|l)$ is a conditional probability density function and $p(l) = \pi_l$ is the prior of the l-th distribution. We model the mapping from joint angles to endpoint forces by a mixture of L Gaussian functions. It is

$$p(d_i|l) = \frac{1}{\sqrt{(2\pi)^D|\Sigma_l|}}\exp(-\frac{1}{2}\left(((\xi_i - \mu_l)^T\Sigma_l^{-1}(\xi_i - \mu_l)\right))$$

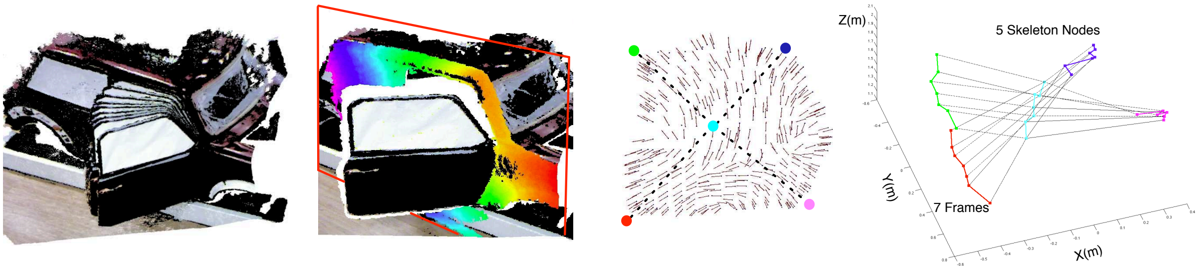
where $\{\pi_l, \mu_l, \Sigma_l\}$ is the Gaussian function's parameter set represented by the prior probability, the mean and covariance matrix. The parameters of the mixture are estimated using the Expectation-Maximization (EM) algorithm. Following learning of the mixture parameters, a generic form of the signals $\boldsymbol{f}_i$ is reconstructed using Gaussian Mixture Regression (GMR). The states s_i are employed as inputs and the output vectors $\hat{f}_i$ are estimated by regression. The mean and covariance matrix of the l-th Gaussian component are defined as

$$\mu_l = \{\mu_{s,l},\ \mu_{f,l}\},\quad \Sigma_l = \begin{pmatrix} \Sigma_{s,l} & \Sigma_{sf,l} \\ \Sigma_{fs,l} & \Sigma_{f,l} \end{pmatrix}.$$

The conditional expectation and covariance of the signal $\boldsymbol{f}_l$ given s are $\hat{f}_l = \mu_{f,l} + \Sigma_{fs,l}(\Sigma_{s,l})^{-1}(s - \mu_{s,l})$, $\hat{\Sigma}_{f,l} = \Sigma_{f,l} - \Sigma_{fs,l}(\Sigma_{s,l})^{-1}\Sigma_{sf,l}$. Finally, the conditional expectation and covariance of $\boldsymbol{f}$ given s for a mixture of K Gaussian components are defined by $\hat{f} = \sum_{l=1}^{L}\beta_l\hat{f}_l$, $\quad\hat{\Sigma}_{\boldsymbol{f}} = \sum_{l=1}^{L}\ \beta_l^2\hat{\Sigma}_{f,l}$, where $\beta_l = p(s|l)/\sum_{j=1}^{L}p(s|j)$ is the responsibility of the l-th Gaussian for s_i. The task force profile $\boldsymbol{f}$ is learnt in the joint space s which is represented by the angle θ.

3 EXPERIMENTAL RESULTS

This paper focuses on skeletonization and manipulat-



a) Data acquirement with different configurations

b) Object extraction based on RANSAC plane estimation

c) Vector field and extracted skeleton topology

d) Skeleton node trace through different frames

Figure 5: Skeletonization of a car door which has a single revolute kinematic joint.

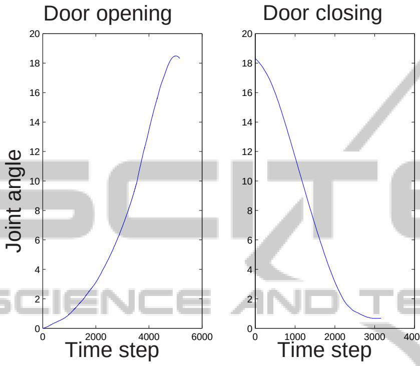


Figure 6: Angle state space estimated based on the position of the car's door handle. The joint angles are expressed in degrees. The time step is equal to 1ms.

ing a single-joint articulated object. We demonstrate the performance of proposed method in a pitstop scenario where the single-joint car door is to be recognized and manipulated. A model of the door, represented by (1), is built which contains the skeleton topology, the kinematic descriptor of the door's joint and the end-effector force required for manipulation.

The point cloud of the door is acquired by one Kinect[1] sensor which is mounted on the top of the robot, shown in Fig. 1. This data is used for skeletonization of the door and estimation of the skeleton node traces over different frames, shown in Fig. 5. The skeletonization of object is realized partially based on the Point Cloud Library [2]. We desire to learn manipulation skills in terms of the force which is required to open or close this single-joint car door.

Appropriate force is demonstrated to the robot by kinesthetic teaching and learned from multiple demonstrations of a task using the proposed approach. Several demonstrations of a door-opening-and-closing task are provided to a 7 DoF robotic arm. Task space force as well as end-effector positional trajectories are captured during demonstrations. Following task space force learning, generalization is required to situations where the initial door position may differ and based on the task goal such as opening or closing. To do so, the force constraints of the task are learnt with respect to door's joint states. The current joint states are estimated using current frame's visual data.

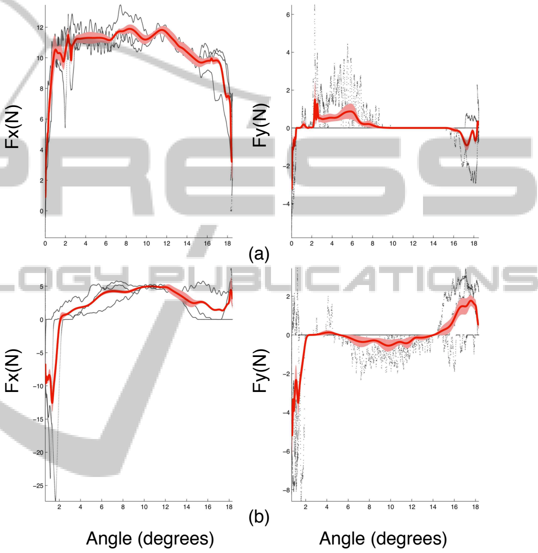


Figure 7: Learning the generalized 2-dimensional force profile of a task in joint angle space given 3 task demonstrations. (a) Door opening, (b) door closing.

Skeletonization of the car door is shown in Fig. 5, where the door is recognized as single-joint articulated object using (3). We observe that the trace of skeleton node has the same motion pattern with the robot arm end-effector trajectory. The current door's joint state could be achieved by the skeleton topology position and learned door's rotational joint model. Every demonstration consists of a door-opening and a door-closing phase without any interruption between the two phases. The different start and end points of each trial are due to slight sliding movement of the robot end-effector along the handle of the door. Given manipulation trajectory, the type of joint is identified firstly by using the algorithm described in 2.2. The

[1] http://www.primesense.com
[2] http://www.pointclouds.org

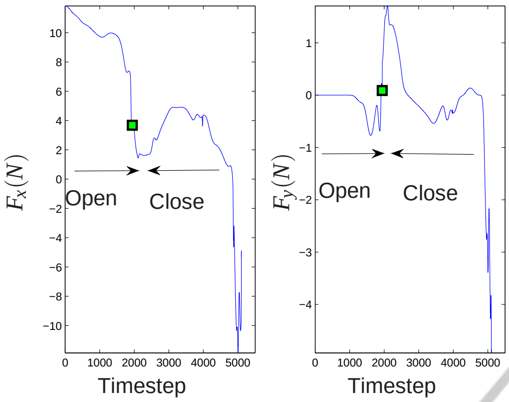


Figure 8: Door opening and closing where the door is initially open at 8 degrees. The time step is equal to 1ms.

door's joint is characterized as revolute and estimate the joint space constrains which is computed, see Fig. 6. This angle space constitutes the input state space in terms of which the force trajectories are learned from multiple demonstrations. Fig. 7 shows learning of the 2-dimensional force for a door opening-closing task from 3 demonstrations by using the method described in Section 2.3. The force is learned separately for the two phases of the task. Following learning, we desire to generalize the force generation policy to different tasks with different current state. More specifically, the case is considered where the car door is already open at 8 degrees and the force profile is estimated which needs to be exerted in order to open the door completely and close it afterwards. Fig. 8 shows the generalized force for this task where the two phases, opening and closing.

4 CONCLUSIONS

In this paper, we propose a method for articulated object modeling by combining visual and haptic data. Visual processing contributes to recognizing the object and identifying its structure and more specifically, its skeleton topology, the number and type of joints as well as the current joint states. Haptic data represented by force are learned from multiple task demonstrations in order to be able to operate the articulated mechanism. The forces are encoded with respect to joint states so that the system can generalize to new situations where the initial object configuration, and thus, joint state differs. The proposed method is demonstrated in manipulation of a single-joint car door. Future work will focus on modeling of a wide-variety of objects which also involve more than one joint.

ACKNOWLEDGEMENTS

This work is supported in part by the DFG excellence initiative research cluster ”Cognition for Technical Systems – CoTeSys,” see www.cotesys.org, the FP7 EU-STREP ”Interactive Urban Robot (IURO)”, see www.iuro-project.eu, the Institute for Advanced Study (IAS), Munich, and the China Scholarship Council (CSC), see http://en.csc.edu.cn.

REFERENCES

Calinon, S., Guenter, F., and Billard, A. (2007). On learning, representing, and generalizing a task in a humanoid robot. *Systems, Man, and Cybernetics, Part B: Cybernetics*, 37(2):286–298.

Huang, X., Walker, L., and Birchfield, S. (2012). Occlusion-aware reconstruction and manipulation of 3d articulated objects. In *In Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 1365–1371, St. Paul, Minnesota.

Kalakrishnan, M., Righetti, L., Pastor, P., and Schaal, S. (2011). Learning force control policies for compliant manipulation. In *Intelligent Robots and Systems (IROS)*, pages 4639–4644.

Katz, D. and Brock, O. (2008). Manipulating articulated objects with interactive perception. In *In Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 272–277, Pasadena, CA.

Lee, D. and Ott, C. (2011). Incremental kinesthetic teaching of motion primitives using the motion refinement tube. *Autonomous Robots*, 31(2):115–131.

Lutscher, E., Lawitzky, M., Cheng, G., and Hirche, S. (2010). A control strategy for operating unknown constrained mechanisms. In *In Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 819–824, Anchorage, Alaska, USA.

Rusu, R. B., Bradski, G., Thibaux, R., and Hsu, J. (2010). Fast 3d recognition and pose using the viewpoint feature histogram. In *In Proc. of the International Conference on Intelligent Robot Systems (IROS)*, pages 2155–2162, Taipei, Taiwan.

Sturm, J., Stachniss, C., and Burgard, W. (2011). A probabilistic framework for learning kinematic models of articulated objects. *Journal of Artificial Intelligence Research*, 41(2):477–526.

Tagliasacchi, A., Zhang, H., and Cohen-Or., D. (2009). Curve skeleton extraction from incomplete point cloud. *In ACM Trans. on Graph*, 28(3):71.

Wang, W., Bršćić, D., He, Z., Hirche, S., and Kühnlenz, K. (2011). Real-time human body motion estimation based on multi-layer laser scans. In *In Proc. of the International Conference on Ubiquitous Robots and Ambient Intelligence*, pages 297–302, Incheon, Korea.