Technical Report

3D Compositional Models associated with Grasp Templates and Haptic Data

Alexander Rietzler ALEXANDER.RIETZLER@UIBK.AC.AT
Institute of Computer Science
University of Innsbruck
Innsbruck, Austria

1. Problem Statement

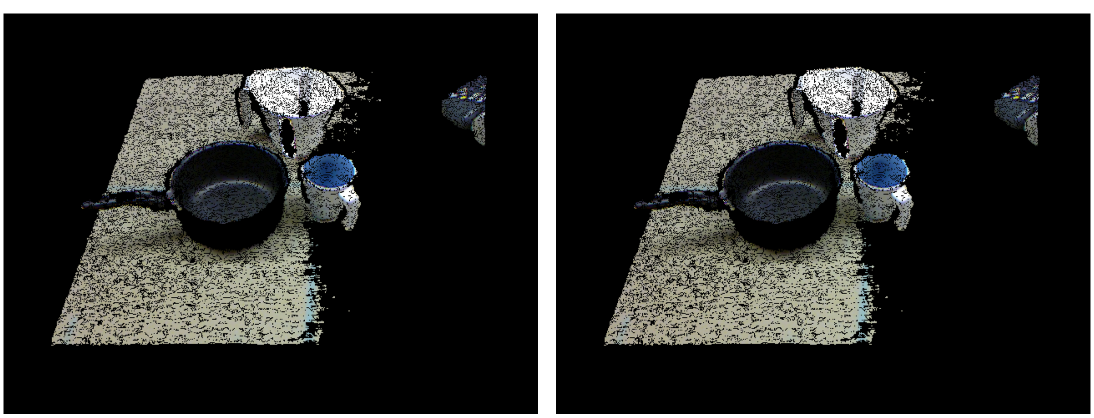
The problem dealt with is to associate graspability and haptic characteristics to parts produced by a 3D compositional models in order to perform robust grasping of even unknown objects exhibiting known parts. The work is strongly tied to requirements of Task 2.1 and 2.2 of the PaCMan project Workpackage 2.

2. Experimental Setup

As a vision sensor, the depth sensor Microsoft Kinect is used. The 3D pointcloud data captured by the sensor is represented as a continuous surface distribution in SE(3) (Detry and Piater, 2010).

In order to perform grasp experiments and to gather tactile data our setup is equipped with a Kuka Lightweight Robot and a Schunk Dextrous Hand 2 with 6 tactile arrays at its 6 finger links.

First a set of grasps are shown on a training set capturing object-relative wrist pose of the gripper, tactile signature and kinesthetic signature (gripper joints), see Figure 1.



(a) training set (b) test set

Figure 1: This figure shows training and test set. To be replaced by other images.

After recognizing meaningful parts of the training objects with the 3D-compositional model (Rezapour-Lakani et al., 2014) the grasps including haptic characteristics are associated to the part the grasp is performed on.

To evaluate the system, a grasp experiment is performed measuring grasp success rates on a test set contained of partly similar objects.

3. Method

Part and Grasp Representation

The method chosen builds on part recognition software developed by Rezapour-Lakani et al. (2014). Given a novel scene, the software segments unseen objects containing similar parts into their known parts. Each part belongs to a previously-learned part cluster marked with a **cluster_id**. To restrict grasp template alignment to the associated part, its oriented bounding box (**OBB**) is computed and passed to grasp inference algorithm.

To each part cluster a set of grasps are attached. In order to match grasps on an unseen but familiar part, the shape-relevant information of graspability is captured by a 3D template $G(x)$, where each point x is represented by a 6D Kernel resulting in a continuous shape distribution. The template is generated by cropping the shape distribution with a cube-shaped region of interest (ROI) close to the gripper contact points during a trained grasp. Grasps are further represented by a pre-grasp pose and a retreat pose, and by 3 grasp types – pinch, spherical, parallel. Figure 2 shows a set of selected templates:

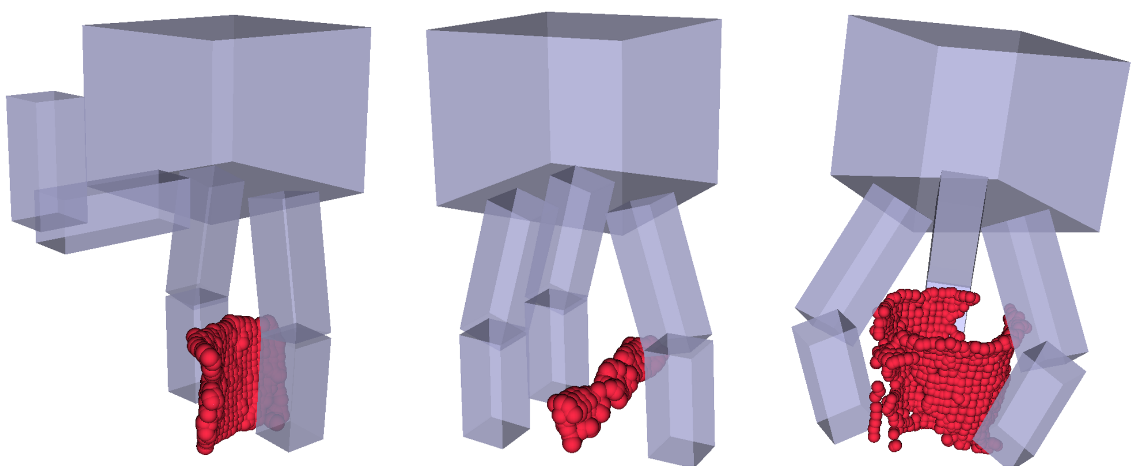


Figure 2: Three selected templates for visualization. From left to right: body-pinch (container), handle-parallel (pot), body-spherical (mug).

Tactile data is gathered during a training grasp. Each of the six phalanges of the gripper is equipped with a tactile array of 6×14 taxels for proximal and 6×13 for distal phalanges, and a sensor value ranging from 0 to 5000 proportional to taxel pressure, here treated as continuous values. For our training grasps only the distal arrays are active, which yields $3 \times (13 \times 6)$ continuous values.

Grasp Inference

The grasp template alignment procedure is based a pose estimation algorithm[1], where the 6D pose-dependent **cross correlation** between the shape distribution of the scene $Q(x)$ and the template $P(x)$ is computed:

$$S(G, Q) = \int G(x) * Q(x)\, \mathrm{d}x \tag{1}$$

This integral is computed by Monte Carlo approximation. In addition, to force the alignment to a selected part only, the OBB of the grasp-template (G) has to intersect with the P of the selected part:

$$C(G, P) = \begin{cases} 1 & OBB(G) \cap OBB(P) \neq \{\} \\ 0 & OBB(G) \cap OBB(P) = \{\} \end{cases} \tag{2}$$

As our grasp inference algorithm relies on optimizing over a large number of hypotheses computing a single best grasp per template, we have to check for collisions with the environment as well as for reachability of the gripper wrist pose hypothesis (r,t) via inverse kinematics during optimization. The reachability term reads

$$R_{r,t} = \begin{cases} 1 & (r, t) \text{ reachable} \\ 0 & (r, t) \text{ not reachable} \end{cases} \tag{3}$$

Combing equations (1–3), our optimization over wrist poses yields

$$S^{*}\mathrm{argmax}_{t,r}\{S(T_{r,t}(G), Q) \cdot C(T_{r,t}(G), P) \cdot R_{r,t})\} \tag{4}$$

where $T_{t,r}(\cdot)$ denotes a rigid body transformation.

Success classification

To be dones:

Here it will be written what features are used based on tactile and haptic data to classify successful and non successful grasps based on:

- haptic data alone
- grasp metric (best S from equation 4)

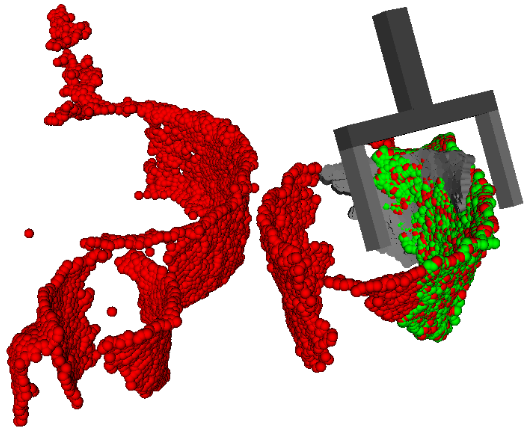
Both classifiers are built per template or grasp type for a small selected number of templates.

4. Experiment

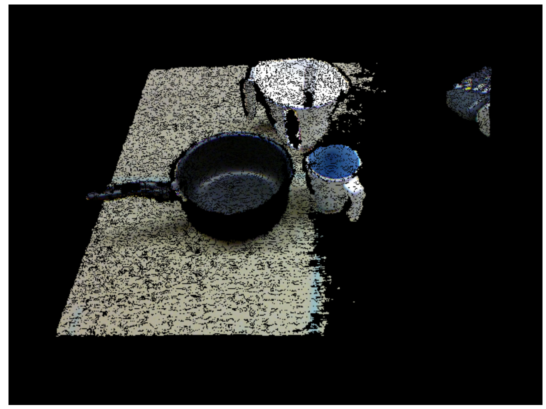
Figure 3 shows a scene with a selected part and corresponding aligned grasp template with highest alignment score.

TODO: Here the grasp success measurement will be described and performance numbers will be given.

1. `http://nuklei.sourceforge.net`



(a) point cloud



(b) RGB image

Figure 3: Instantiated scene containing 3 objects. Points in green shows the selected part the grasp template is aligned to. Points in grey show the aligned grasp template. The gripper indicates wrist pose and grasp type. Here a pinch grasp has the highest alignment score.

References

Renaud Detry and Justus Piater. Continuous surface-point distributions for 3D object pose estimation and recognition. In Ron Kimmel, Reinhard Klette, and Akihiro Sugimoto, editors, *Asian Conference on Computer Vision*, volume 6494 of *LNCS*, pages 572–585, Heidelberg, 2010. Springer. doi: 10.1007/978-3-642-19318-7_45. URL `https://iis.uibk.ac.at/public/papers/Detry-2010-ACCV.pdf`.

Safoura Rezapour-Lakani, Mirela Popa, Antonio J. Rodríguez-Sánchez, and Justus Piater. Scale-Invariant, Unsupervised Part Decomposition of 3D Objects. In *Parts and Attributes*, 9 2014. URL `https://iis.uibk.ac.at/public/papers/Rezapour-2014-PA.pdf`. Workshop at ECCV.