One shot contact learning

Marek Kopicki and Jeremy L. Wyatt
School of Computer Science, University of Birmingham, UK
Email: {msk,jlw}@cs.bham.ac.uk

***Abstract*—We present a general method for one-shot learning of dexterous grasps and object placement. We rely on a product of experts formulation. Experts in the product are of two types. The first is a *contact model* and is a density over the pose of a single hand link relative to the local object surface. The second is the *hand configuration model* and is a density over the whole hand configuration. Grasp generation for a novel object optimises the product of these two model types. Selection between grasp types is also automatic. The success rate is 84.4% or 77.7% if seven views or a single view of the test object are taken. We also show initial results on object placement.**

Fig. 1: **Left:** The four objects on the left were used for training, the forty three objects on the right were used as test objects. **Right:** The Boris manipulation platform.

I. Introduction

Previous work in learning generalisable grasps falls broadly into two classes. One class utilises the shape of common object parts or their appearance to generalise grasps across object categories [8, 2, 3, 7]. This works well for low DoF hands. Another class captures the global properties of the hand shape [1]. This global hand shape can be associated with global object shape, allowing generalisation by warping grasps to match warps of global object shape [4]. This second class works well for high DoF hands, but generalisation is more limited. We achieve the advantages of both classes, generalising grasps across object categories with high DoF hands. The method can be simply adapted to learn object placement in the same manner, here we show learning of placing different sized plates into a dishrack. Our grasping work is described in more detail in [5, 6].

II. Approach

The technical innovation to achieve this is to learn two types of models from the example grasp (Figure 3 left and centre), and then recombine them using a product of experts formulation when inferring a new grasp (Figure 3 right). Dexterous grasping involves simultaneously satisfying multiple constraints, and our central insight is that a product of experts is a natural way to encode these. Both model types are density functions. The first is a *contact model* of the relation between a rigid link of the hand, and the local object shape near its point of contact (Figure 3 left). We learn one contact model for each link of the hand involved in the grasp, and these capture local constraints in the grasp. To capture global information we learn a second type of model, a *hand configuration model* from the example grasp (Figure 3 centre).

When presented with a novel object the contact model is combined with the point cloud of the new object to create a contact query density. Figure 2 shows the contact query densities for two finger links, where the training grasp was on a bowl and the test object is a kettle. It can be seen that the finger links (marked in blue) generate different densities over the kettle. A contact query density is calculated for the new object point cloud, for each finger link in each trained grasp. This uses a Monte Carlo procedure which is fast (¡1 sec for four training grasps).

Then grasp generation is performed, followed by grasp improvement (Figure 3 right). The grasp generation proceeds as follows: a finger link is picked at random, a pose on the new object is sampled from its contact query density, and a hand configuration is sampled from the hand configuration model for the grasp type. By forward kinematics this defines a complete hand pose. Thus the hand configuration model constrains the combined search space for the link placements. Many such grasps are generated on the new object for each grasp type. In the final stage these grasps are optimised. Grasp improvement is carried out by performing simulated annealing on a product of experts expressing the grasp likelihood for each candidate. The optimised grasps are ranked by this likelihood. The whole process takes 5-25 secs

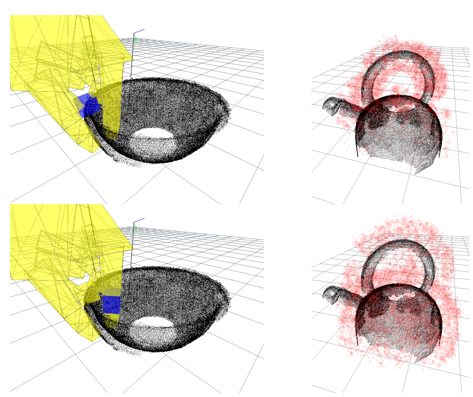

Fig. 2: Contact query density (red cloud) for two links (blue).

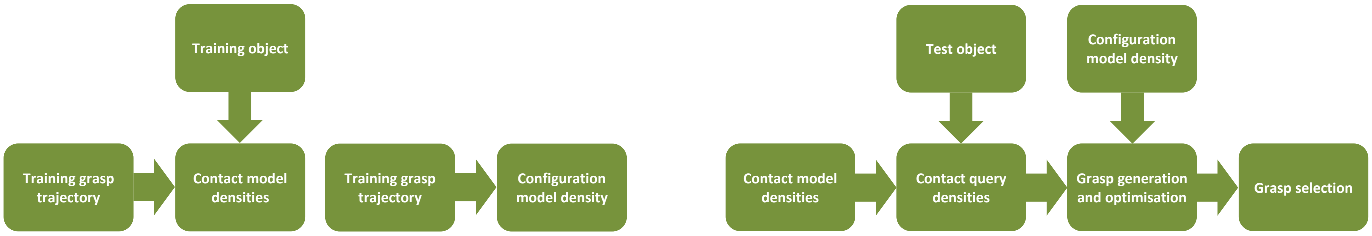

Fig. 3: Grasp training (left) and testing procedure (right).

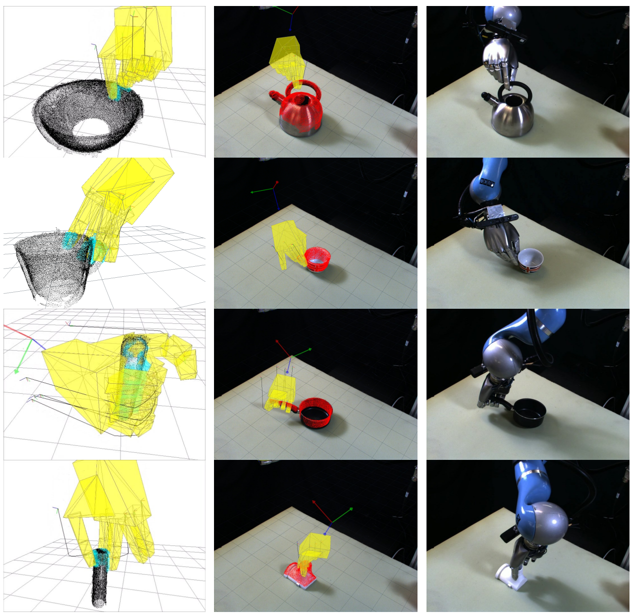

Fig. 4: Grasp transfer. Training grasp in left column. Grasp for test object in the centre and right columns.

on a standard PC, where between 500 and 2500 grasps are generated and optimised. The ranking of the individual grasps by likelihood can be seen as a probabilistic, factored, memory based grasping. Once normalistion is performed then grasps generalised from different training grasps are interleaved in this ranking, and so the method automatically enables simultaneous selection and adaptation of grasp type. After grasp optimisation and ranking, kinematic infeasible grasps are discarded, and the top ranked grasp is executed. Given greater reconstruction of the test object surface the success rate is higher, so that seven views from a depth camera give a success rate of 84.4%, and one view gives a success rate of 77.7%. The generalisation ability is shown in Figure 4. The method is also applicable to learning object placement. Here the placement is shown for test plates, of varying sizes and shapes. Placement is into a dishrack, which acts as the receiving 'hand' and the learning is able to infer various poses and approaches for the object to be stably placed in the dishrack. Figure 5 shows two test plates, with their visualised positions. This shows that the same approach of products of experts is able to learn to generalise both grasping and placement from one training example.

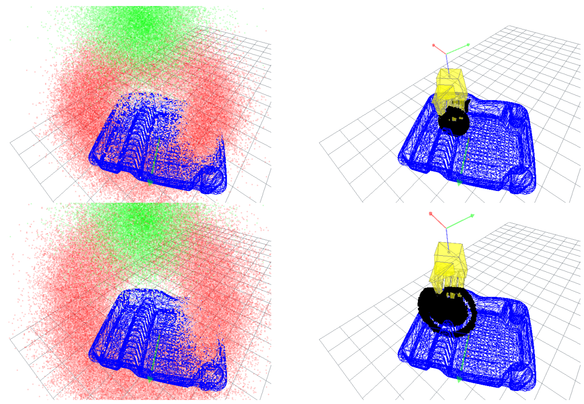

Fig. 5: Small and large plate dishwasher placement example. Left: wrist pose densities for the dishwasher-plate contact (red) and the plate insertion trajectory (green). Right: computed maximum likelihood wrist pose of the densities product.

The authors would like to thank the support of EC project PaCMan-FP7-IST-600918.

References

[1] H. Ben Amor, O. Kroemer, U. Hillenbrand, G. Neumann, and J. Peters. Generalization of human grasping for multi-fingered robot hands. In *International Conference on Intelligent Robots and Systems*, pages 2043–2050. IEEE, 2012.

[2] R. Detry, C. Henrik Ek, M. Madry, and D. Kragic. Learning a dictionary of prototypical grasp-predicting parts from grasping experience. In *IEEE ICRA*, pages 601–608. IEEE, 2013.

[3] Alexander Herzog, Peter Pastor, Mrinal Kalakrishnan, Ludovic Righetti, Jeannette Bohg, Tamim Asfour, and Stefan Schaal. Learning of grasp selection based on shape-templates. *Autonomous Robots*, 36(1-2):51–65, 2014.

[4] U. Hillenbrand and M.A. Roa. Transferring functional grasps through contact warping and local replanning. In *IEEE/RSJ International Conference on Robotics and Systems*, pages 2963–2970. IEEE, 2012.

[5] Marek Kopicki, Renaud Detry, Florian Schmidt, Christoph Borst, Rustam Stolkin, and Jeremy L. Wyatt. Learning dextrous grasps that generalise to novel objects by combining hand and contact models. In *IEEE ICRA*, pages 5358–5365. IEEE, 2014.

[6] Marek Kopicki, Renaud Detry, Maxime Adjigble, Rustam Stolkin, Ales Leonardis, and Jeremy L. Wyatt. One shot learning and generation of dexterous grasps for novel objects. *IJRR*, 2015.

[7] O. Kroemer, E. Ugur, E. Oztop, and J. Peters. A kernel-based approach to direct action perception. In *IEEE ICRA*, pages 2605–2610. IEEE, 2012.

[8] A. Saxena, L. Wong, and A.Y. Ng. Learning grasp strategies with partial shape information. In *Proceedings of AAAI*, pages 1491–1494. AAAI, 2008.