Grasp Moduli Spaces, Gaussian Processes, and Multimodal Sensor Data

Florian T. Pokorny, Yasemin Bekiroglu, Johannes Exner, Mårten Björkman, and Danica Kragic
Centre for Autonomous Systems/CVAP, CSC
KTH Royal Institute of Technology
Stockholm, Sweden

***Abstract*—We discuss our recent and ongoing work on deformation-based grasp synthesis and transfer which formulates a joint space of shapes and grasps – a Grasp Moduli Space – within which both grasp configurations and object shapes can be continuously deformed. In this context, we propose the use of Gaussian Process-based implicit surface representations. These shape representations complement our previous work on using spherical harmonics as well as explicit cylindrical coordinates to model the object shape component of a Grasp Moduli Space. We provide a preliminary experiment with these Gaussian Processes showing how shape representations can be obtained using multimodal visual and haptic information and discuss how these representations can be continuously deformed for the purpose of transferring and generalizing known grasps.**

I. Introduction and Motivation

Humans are able to grasp and manipulate objects seemingly with little effort. Furthermore, we are able to easily *generalize* grasps and manipulation sequences from a specific demonstration, and we can *transfer* grasps between similar shapes such as bottles or tools. In robotics, on the other hand, the predominant approach to grasp synthesis has been to solve the grasping problem for each specific object instance by completely reinitializing the developed algorithms at hand. Popular approaches to grasp synthesis such as the sampling-based approach of GraspIT [7], various heuristics-based approaches such as a reduction of the object model onto its medial axis [11], and approaches decomposing the object into parts such as boxes [6] have been developed and typically require a new execution of the underlying grasp algorithm when a new object is encountered.

When we think about the graspable objects surrounding us, we observe that one can consider almost any type of object as being part of *family of similar objects* which vary *continuously* with respect to various parameters such as, most trivially, height, width and depth, but also in certain latent style parameters. Handles of tools such as screw drivers can vary in their precise shape, for example, while it is reasonable to assume that such shapes should be concentrated around a mean shape resembling a cylinder. Humans are able to quickly adapt to such variations - a capability which we aim to endow robots with in this ongoing work. Our approach to grasp transfer is based on a continuous deformation of a known grasp and shape to novel objects with similar shape.

The key representation which we aim to develop further for this purpose is the notion of a *Grasp Moduli Space* [9] modeling shapes and grasps jointly. The term moduli space here is inspired by its use in deformation theory in mathematics [1]. To represent shapes, we are investigating Gaussian Process implicit surfaces [13, 4, 12] providing a probabilistic shape estimate from point-cloud data.

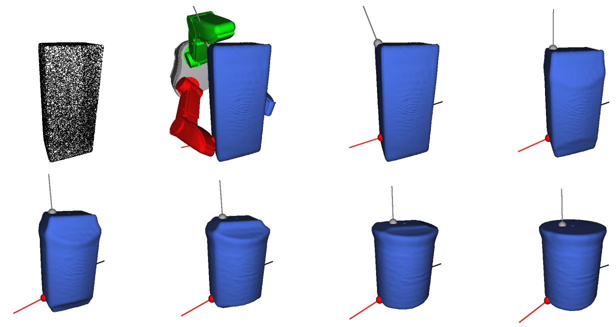

Fig. 1: A joint grasp/shape deformation and optimization in $\mathcal{G}^{sph}$ (see [10]).

II. Grasp Moduli Spaces

In order to model both a shape and a particular grasp on a shape, we need to firstly define an appropriate *shape space* modeling the admissible objects. Secondly, we need to make precise what we mean by a grasp. Here, we consider a grasp g as a tuple of m contact points $c_i \in \mathbb{R}^3$ on an object's surface, unit normal vectors $n_i \in \mathbb{S}^2$ at those contact points, and a center of mass $z \in \mathbb{R}^3$: $g = (c_1, \ldots, c_m, n_1, \ldots, n_m, z) \in \mathbb{R}^{3m} \times (\mathbb{S}^2)^m \times \mathbb{R}^3$. In order to determine whether a grasp is force-closed or not [2], we consider the L^1 grasp quality function Q defined by [5]. Q, for m contact points, is then a function on the $\mathbb{R}^{3m} \times (\mathbb{S}^2)^m \times \mathbb{R}^3$, and this function is in fact tame in the sense that it is Lipschitz continuous with a particular, easily computable, Lipschitz constant under variations in grasps, as shown in our recent work [8]. Grasps with positive grasp quality under Q are then considered more stable the larger Q is. We have recently begun to consider an appropriate simple space of smooth surfaces which allows for a deformation-based grasp synthesis [9]. There, we considered the moduli space of smooth surfaces with cylindrical coordinates and with grasps with m contact positions – called $\mathcal{G}^{cyl}$ – as well as the space of surfaces of revolution with grasp configurations with m contact positions $\mathcal{G}^{rev} \subset \mathcal{G}^{cyl}$. Both $\mathcal{G}^{rev}$ and $\mathcal{G}^{cyl}$ are infinite dimensional and the underlying shape spaces of surfaces of revolution $\mathcal{M}^{rev}$ and of surfaces with cylindrical coordinates $\mathcal{M}^{cyl}$ are convex allowing us to construct shape space subsets from finitely many example surfaces by means of convex combinations of those surfaces.

Similarly, we can jointly and continuously deform a grasp configuration in cylindrical coordinates *while the underlying object shape is deformed*. In [9], we combined such continuous deformations in grasps and shapes, where an initial stable grasp on a surface was continuously optimized using a gradient ascent while the underlying surface was deformed towards a novel shape. Recently [10], we have further extended this framework to be applicable directly with point cloud-data by means of a shape space defined using spherical harmonics. There, we defined a shape space of surfaces with spherical coordinates $\mathcal{M}^{sph}$ and a resulting Grasp Moduli Space $\mathcal{G}^{sph}$ of grasp configurations on such surfaces. Fig. 1 illustrates an example grasp/surface deformation in $\mathcal{G}^{sph}$ with 3 contact points. An interesting aspect of this approach is that it allows us to quantify uncertainty in grasping both in terms of *shape uncertainty*, for example due to partial or unreliable sensor data, and *contact point uncertainty*, *e.g.* when we cannot precisely position the robot hand.

III. Gaussian Processes and Shape Deformations

A natural framework within which noisy data can be described is that of Gaussian Process Regression. We are currently investigating two specific kernel choices to allow us to represent shapes and their deformations in a continuous manner. In the recent work [3], we have shown that the thin plate kernel can provide useful shape representations from a single view kinect point-cloud P if this point-cloud is supplemented by a set P' of points obtained from haptic data of a Schunk Dexterous Hand touching the object of interest, and where an exploration strategy is guided towards points of maximal variance under the GP regression model. There, an object's surface is represented as an implicit surface $S = f^{-1}(0)$, and $f : \mathbb{R}^3 \to \mathbb{R}$ is described using the thin-plate kernel $k(\mathbf{x}_i, \mathbf{x}_j) = 2|r|^3 - 3Rr^2 + R^3$, where $r = |\mathbf{x}_i - \mathbf{x}_j|$ and we set $R = \sqrt{3}L$, and L is the kinect data's bounding box side-length (see [3]). Yet another interesting GP kernel is the Matérn kernel which reduces to the exponential kernel for its parameter $\nu = 0.5$ and to the Gaussian kernel as $\nu \to \infty$. We are in particular considering the Matérn kernel $k_{\nu=\frac{3}{2}}(r) = (1 + \frac{\sqrt{3}r}{l}) \exp(-\frac{\sqrt{3}r}{l})$ with length-scale l.

An advantage of representing a shape by a GP is that we can obtain an estimate for the variance as well as the mean shape determined by the sensor data. Observe that, since a shape is represented by a mean function m. We can, for any two shapes S_1, S_2 corresponding to GP means m_1, m_2 consider a curve $\gamma(t) = (1-t)m_1 + tm_2$, $t \in [0, 1]$ deforming S_1 to S_2. Since Gaussian Processes can model a rich family of shapes, we propose to study these in the context of the Grasp Moduli Space framework. In Fig. 2, we provide an initial illustration of an experiment where two GP-based shape representations are being deformed into each other using the above convex combination of their means. The initial and final shapes are determined from multimodal sensor data from [3], with red points corresponding to single view visual data from a Kinect sensor, while blue points were obtained from haptic sensors on a Schunk Dexterous Hand. Since these GPs represent a shape not as a parametrized surface as in [9, 10], but only as a level set, the representation of grasp contact points in this setting will have to be adapted. A natural Grasp Moduli Space $\mathcal{G}$ consists of mean functions corresponding to smooth embedded surfaces and grasp contact points on the surfaces S_h parametrized by h. More generally, this could be extended to include the full robot hand configuration during a grasp. Many open problems, which we are looking forward to discuss during this workshop, exist in this direction:

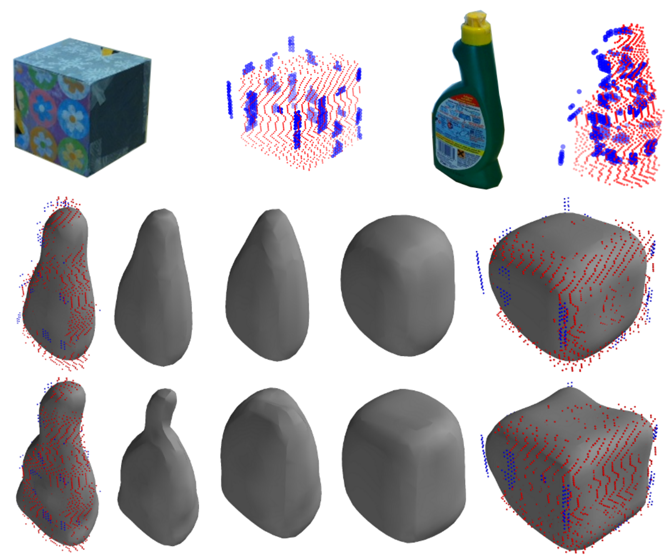

Fig. 2: Box and spray bottle models with haptic (blue points) and sub-sampled single view visual (kinect, red points) sensor data are displayed in the first row. Shape deformations $\gamma(t)$ between these shapes based on a Matérn kernel GP (middle row) and a thin plate GP (bottom row) are displayed for t equal to 0, 0.06, 0.12, 0.5, 1.0 from left to right. Data points are assigned function value 0 and are supplemented by an interior point (value -1) and 15 exterior points (value 1). We transformed the point-cloud data to lie in a cube centered at the origin and of side-length $L = 3$ and set $R = \sqrt{3}L$ for the thin-plate kernel and $l = 1$ for the Matérn kernel. In our Gaussian Process regression, we also assume that our sensor data is corrupted by a Gaussian noise term with $\sigma^2 = 0.1$.

- Which optimization methods are most effective in optimizing a grasp's quality, and more generally a grasp's *task specific utility*, in a general Grasp Moduli Space where shapes are parametrized using Gaussian Processes?
- How can the GP's variance information be incorporated in a deformation-based grasp synthesis framework?
- How can grasp configurations best be modeled probabilistically in conjunction with the GP's shape estimate?
- How can prototypical grasp/shape configurations in $\mathcal{G}$ be determined automatically?

As a first step, we are currently experimenting with various optimization methods which optimize the robot hand's joint configuration while we deform a known shape/grasp configuration towards a novel shape. Methods beyond a simple gradient ascent will likely be necessary here to ensure a rapid convergence and the numerical stability of the optimization of the hand's joint configuration while maintaining contact during a shape and grasp deformation.

References

[1] D. D. Ben-Zvi. Moduli spaces. In Timothy Gowers, June Barrow-Green, and Imre Leader, editors, *The Princeton companion to mathematics*. Princeton Univ. Press, 2009.

[2] A. Bicchi and V. Kumar. Robotic grasping and contact: A review. In *IEEE ICRA*, pages 348–353 vol.1, 2000.

[3] M. Björkman, Y. Bekiroglu, V. Högman, and D. Kragic. Enhancing visual perception of shape through tactile glances. In *IEEE/RSJ IROS*, pages 3180–3186, 2013.

[4] S. Dragiev, M. Toussaint, and M. Gienger. Gaussian process implicit surfaces for shape estimation and grasping. In *IEEE ICRA*, pages 2845–2850. IEEE, 2011.

[5] C. Ferrari and J. Canny. Planning optimal grasps. *IEEE ICRA*, pages 2290–2295, 1992.

[6] K. Huebner, S. Ruthotto, and D. Kragic. Minimum volume bounding box decomposition for shape approximation in robot grasping. In *IEEE ICRA*, pages 1628–1633, 2008.

[7] A. T. Miller and P. K. Allen. Graspit!: A versatile simulator for robotic grasping. *IEEE Robotics and Automation Magazine*, 11:110–122, 2004.

[8] F. T. Pokorny and D. Kragic. Classical grasp quality evaluation: New theory and algorithms. In *IEEE/RSJ IROS*, 2013.

[9] F. T. Pokorny, K. Hang, and D. Kragic. Grasp moduli spaces. In *Proc. of Robotics: Science and Systems*, 2013.

[10] F. T. Pokorny, Y. Bekiroglu, and D. Kragic. Grasp moduli spaces and spherical harmonics. In *IEEE ICRA*, 2014.

[11] M. Przybylski, T. Asfour, and R. Dillmann. Planning grasps for robotic hands using a novel object representation based on the medial axis transform. In *IEEE/RSJ IROS*, pages 1781–1788, 2011.

[12] C. Walder, B. Schölkopf, and O. Chapelle. Implicit surface modelling with a globally regularised basis of compact support. In *Computer Graphics Forum*, volume 25, pages 635–644. Wiley Online Library, 2006.

[13] O. Williams and A. Fitzgibbon. Gaussian process implicit surfaces. *Gaussian Proc. in Practice*, 2007.