Grasping by interconnection: robust manipulation with minimal object information

Julien Vanderheyden[1], Guillaume Drion[1], Pierre Sacré[1,*], and Fulvio Forni[2,*]

***Abstract*— Dexterous grasping in unstructured environments remains a central open challenge in robotics, requiring simultaneous robustness to perceptual uncertainty and generalization across diverse object geometries. Existing approaches rely on accurate object models, either predefined or reconstructed online, making them brittle to the sensor noise inherent in real-world settings. We propose a grasp planning strategy that couples the robotic hand to minimal primitive-based object representations. Through virtual mechanical interconnections, each primitive shape corresponds to a grasp type and generates fully adaptive closing trajectories. Experimental evaluation across a wide set of objects demonstrates robustness to pose estimation errors of several centimeters in translation and several tens of degrees in orientation. This work demonstrates that robust, task-oriented dexterous grasping can emerge from minimal object information, opening a path toward reliable manipulation in unstructured real-world scenarios.**

I. Introduction

Dexterous grasping in unstructured environments remains an open challenge in robotics [1]. The high-bandwidth feedback of distributed, discontinuous contacts between the hand and the object requires significant design efforts in both mechanical hardware and control algorithms. The complexity of hand-object interaction often leads to failures, which can be mitigated through detailed information about the target object alongside reliable, accurate sensing. In practice, however, given the vast variety of graspable objects, details are often uncertain, and visuo-tactile sensing is hampered by significant sensor noise, including uncertain nonlinearities and occlusions. Two objectives are therefore central: achieving robustness to perceptual uncertainty and ensuring generalization across diverse objects.

Current approaches to grasping typically rely on a precise object model, obtained either from prior knowledge or from sensor measurements. However, relying on a library of precise object models limits generalization. Mitigations through sensor-based reconstruction introduce their own challenges. Perception is often compromised by visual clutter, occlusions, and changing lighting conditions, which lead to noisy or incomplete object estimates. In this paper, we advocate the use of *shape primitives* as object representations that remain stable under perceptual uncertainty while enabling sufficient expressiveness for adaptive grasping.

This work was supported by the Belgian Government through the Federal Public Service Policy and Support.

[1]J. Vanderheyden, G. Drion, and P. Sacré are with the Department of Electrical Engineering and Computer Science, University of Liège, Belgium (julien.vanderheyden@uliege.be; gdrion@uliege.be; p.sacre@uliege.be).

[2]F. Forni is with the Department of Engineering, University of Cambridge, United Kingdom (f.forni@eng.cam.ac.uk).

[*]PS and FF contributed equally to this work as senior authors.

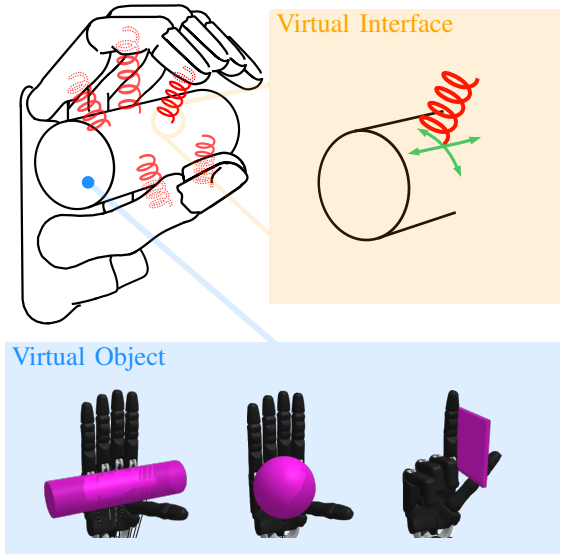

Fig. 1. Robust grasping emerges from minimal information. Each shape primitive is embodied as a *virtual object*; the hand is coupled to this virtual object through a *virtual interface* driven by virtual model control (VMC), generating fully adaptive finger motion without a precise object model.

Shape primitives, such as cylinders, spheres, and flat boxes, provide minimal object abstractions, which are easier to extract from limited sensing information (*e.g.*, point-cloud data). We pair them with canonical grasp types [2], so that both perception and action are organized around a small set of discrete interaction patterns. Hand motion is then generated through a *virtual dynamic mechanical interface* between the robot and the primitive, taking advantage of virtual model control (VMC) [3]–[6]. As illustrated in Fig. 1, the virtual mechanical elements of the interface *connect* robot hand and shape primitives to generate virtual *forces* that drive the fingers motion. Virtual model control produces adaptive grasping without requiring a precise object model or predefined trajectories, making it well suited to the uncertainty inherent in unstructured environments.

Our grasp planning strategy is experimentally evaluated on the Shadow Dexterous Hand [7]. Results demonstrate significant robustness and generalization across a wide range of object shapes and dimensions. The primary contributions of this paper are: (i) a novel grasp planning strategy for robust manipulation with minimal object information, leveraging the virtual model control framework, and (ii) an experimental evaluation on the Shadow Dexterous Hand platform, validating the system performance across a diverse set of object types, scales, and geometries.

II. Related work

A. Standard grasping approaches

Analytical and data-driven grasping methods achieve strong performance under controlled conditions, but show sensitivity to object model quality that limits their applicability in uncertain environments.

Analytical methods address the challenge of computing force-closure grasps by optimizing over contact geometry [8], [9]. However, these techniques typically require detailed object models, which are either assumed to be known *a priori* or must be reconstructed online from partial observations. This requirement often limits generalization to a predefined set of objects [10]. Furthermore, online reconstruction from partial data frequently introduces robustness issues due to the noise inherent in real-world 3D sensing [11]. In both cases, the reliability of the grasp plan is constrained by the "bottleneck" of object representation and reconstruction accuracy.

Data-driven approaches learn grasp strategies from large datasets of successful grasps, implicitly constructing internal object representations. Generalization to unseen objects can be achieved either by reconstructing the object shape explicitly [12] or by direct inference from raw point clouds [13]. Although effective, both strategies implicitly rely on the assumption that the input data quality at test time matches the training distribution, making them sensitive to sensor noise and domain shift. Furthermore, acquiring the large amounts of high-quality data for training is costly and difficult to scale to novel object categories.

B. Object shape approximation

An alternative to precise object reconstruction is to represent the object using simple geometric primitives. The principle is to reduce dependency on geometric detail in exchange for increased robustness to degraded sensing [14]. This choice is further supported by evidence from cognitive science, which shows that human object manipulation depends more on global shape characteristics than on precise geometric detail [15], [16], suggesting that primitive-based representations align closely with natural grasping.

Among the simplest strategies, generic shape primitives, such as cylinders, spheres, and boxes [14], [17], and box-based representations [18]–[20] have been particularly successful in finding grasp regions. These methods focus on predicting a target hand pose. Each finger trajectory is determined at planning time and executed. As a consequence, they cannot adapt the grasp to the geometric variation of objects within the same primitive category. Advanced strategies explore more granular object representations. These are based on richer representations such as superquadrics [21], [22], medial axes [23], [24], and skeletal models [25]. Although these enable more accurate motion planning, their dependence on accurate shape estimation makes them more vulnerable to sensor noise.

This trade-off between geometric expressiveness and robustness motivates the use of simpler primitives in this work. Our work embraces the simplicity of shape primitives while addressing their limitations through virtual model control. VMC regulates the finger motion in closed loop, through a (virtual) *compliant* robot-object interaction layer. The result is that each shape primitive determines the motion features of the fingers, but their real-time motion remains capable of conforming to the real object geometry encountered at execution time.

C. Human-inspired grasp synthesis

Taking inspiration from human behavior provides a principled way to manage the high dimensionality of the grasping problem. *Grasp taxonomies* organize grasps into canonical types, providing lightweight structural priors that guide hand configuration without solving a full high-dimensional optimization problem. The effectiveness of this approach has been confirmed empirically. Explicit models of grasp type significantly improve task success, even in data-driven settings [26]. However, early approaches paired each grasp type to a fixed closing trajectory [17], [27]–[29], reducing adaptability to the object geometric variations. Methods that combine grasp type and analytical optimization [25], [30]–[34] take advantage of the narrower search space, but still require a detailed object model to evaluate grasp quality during optimization. Likewise, methods that rely on grasp type in learned policies [35]–[38] reduce geometric dependency but require large datasets, as discussed in Section II-A. In summary, the advantage of the lightweight priors offered by grasp taxonomies appears to be often hindered by rigid or model-dependent executions.

Hand synergies offer a powerful complementary approach to reduce the complexity of hand control. The core idea is to exploit the low-dimensional structure of natural hand motion. Postural synergies [39], [40] capture the dominant modes of joint covariation observed in human grasping. They have been used to define a compact basis of hand postures sufficient for a wide variety of grasps, making both optimization and learned control tractable in real time [35], [41]. However, synergies are typically defined independently of the object being grasped: the same basis is applied regardless of object shape, and not all hand configurations are reachable within the subspace, which limits dexterity. In addition, because the subspace is fixed by design, the hand configuration cannot adapt to geometric variation encountered during execution.

In this paper we take advantage of grasp types to identify continuous regions of high-quality configurations, which are later refined by, or adapted to, the variations in object size and geometry. Rather than constraining the hand to a predefined low-dimensional synergy subspace, we propose virtual model control as an alternative for navigating within these grasp regions. Through structured, low-dimensional, mechanical coupling between the robot and the object primitive, VMC generates coordinated and adaptive hand motion without sacrificing kinematic expressiveness. The taxonomy determines which coupling is appropriate for a given primitive; the VMC dynamics then handles fingers coordination.

III. VMC FOR GRASPING

A. Design approach

A virtual model controller is a passive system composed of virtual mechanical elements (springs, dampers and linkages), virtually connected to the controlled robot [4], [5]. The associated (simulated) forces determine the robot motion. In the classical case of robot manipulators with revolute joints, these virtual forces $\boldsymbol{F}_i$ act on generic (interconnection) points $\boldsymbol{z}_i = h_{\boldsymbol{z}_i}(\boldsymbol{q})$, here expressed as functions of the joint coordinates $\boldsymbol{q}$. These forces are translated into joint torques by

$$\boldsymbol{u} = \sum_{i=1}^{n} J_{\boldsymbol{z}_i}^T(\boldsymbol{q})\, \boldsymbol{F}_i, \tag{1}$$

where $J_{\boldsymbol{z}_i} = \frac{\partial h_{\boldsymbol{z}_i}(\boldsymbol{q})}{\partial \boldsymbol{q}}$ is the corresponding Jacobian matrix.

Virtual model control guarantees that the controlled robot remains passive [42]–[45] (within the limits of faithful approximation of the virtual mechanical elements). This guarantees stable interaction with passive objects without requiring precise contact models, a critical advantage in handling stiff contact transitions. Virtual model control also avoids the heavy inverse kinematics that the high dimensionality of dexterous hands would otherwise require. Finally, the representation of a controller as a virtual mechanism provides a modular architecture, allowing for grasp-specific elements to be added or reconfigured independently. It makes it straightforward to pair controller with shape primitives and to handle multiple grasp types within a common framework.

Our approach frames grasping as a mechanical interconnection between the robotic hand and the simplified object model (the shape primitive). Our motion planning strategy is based on two key components: a *virtual object*, which models the shape primitive through VMC elements, and a *virtual interface*, which couples the hand to the virtual object. Three primitive object shapes are considered: a cylinder, a sphere, and a flat box. These are related to the three grasps: medium wrap, power sphere, and lateral pinch, identified by Bullock et al. [46] as the most versatile ones in household settings. This pairing provides compact and complete coverage of common object categories with a minimal set of canonical interaction patterns.

B. Virtual object

Virtual objects generate repulsive forces that drive the fingers toward stable contact configurations, avoiding finger–object interpenetration.They are constructed from a base shape (cylinder, sphere, flat box) with ReLU-shaped springs of the form

$$\boldsymbol{F}_{\mathrm{s}} = k_{\mathrm{s}} \min\left(|\boldsymbol{z}| - z_{\mathrm{rest}}, 0\right) \frac{\boldsymbol{z}}{|\boldsymbol{z}|}, \tag{2}$$

where $\boldsymbol{z}$, z_{rest}, and k_{s} are the extension, rest length, and stiffness of the spring, respectively. The spring generates a repulsive force once a compression threshold is exceeded, thereby mimicking the onset of contact.

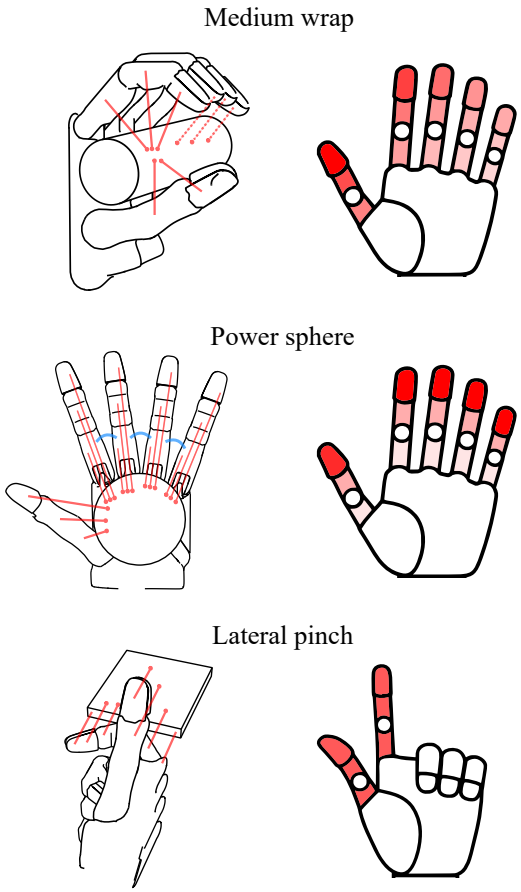

Fig. 2. Each grasp type is encoded by a distinct virtual interface architecture and stiffness distribution. *Left.* Contractive (red) and repulsive (blue) springs drive the fingers toward stable contact; sliding terminals (circles) adapt continuously to object size. *Right.* Higher color intensity indicates higher stiffness.

Dampers are introduced to suppress oscillations at contact. The associated forces read $\boldsymbol{F}_{\mathrm{d}} = c(|\boldsymbol{z}|)\, \dot{\boldsymbol{z}}$ with ReLU-shaped damping

$$c\left(|\boldsymbol{z}|\right) = \frac{c_{\max}}{\gamma z_{\mathrm{rest}}} \max\left(\gamma z_{\mathrm{rest}} - |\boldsymbol{z}|, 0\right). \tag{3}$$

The coefficient γ ($\gamma = 1.05$) slightly inflates the virtual object so that damping activates before the spring force takes effect, ensuring a smooth behavior. Continuous force profiles are used throughout to avoid abrupt variations and ensure smooth resulting motions. In general, a wider set of passive contact models could be considered without altering the overall architecture.

C. Virtual interface

Grasp synthesis inherently involves two complementary design problems: identifying a mechanically stable final contact configuration, and shaping a robust transient motion that brings the hand to that configuration. In addition, different grasp types impose specific constraints that must be explicitly encoded. The virtual interface addresses these three aspects sequentially, as shown in Fig. 2 and detailed below.

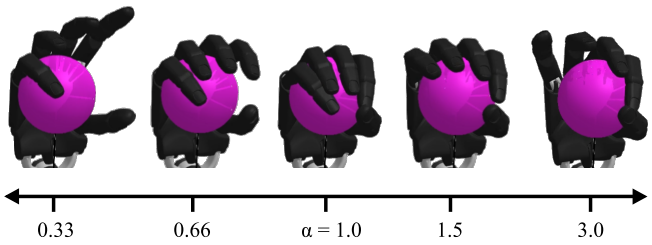

Fig. 3. A single parameter governs the coordination between fingers during closure. At $\alpha = 1$, all fingers close symmetrically. Increasing α advances the thumb and index first, creating a stabilizing pinch before full-hand enclosure; decreasing α advances the outer fingers first.

Stable contacts are achieved by connecting virtual springs between the virtual object and specific points on the hand, with the goal of driving the fingers and palm toward the object. These connection points can slide along the object surface via prismatic and revolute joints, enabling adaptation to varying object dimensions. The hand configuration adapts continuously with object size without requiring re-planning.

The *transient motion* of the finger towards the stable contact configuration outlined above is regulated through placement and tuning of multiple virtual springs. We consider two complementary mechanisms, both inspired by characteristic features of human motor behavior.

The first mechanism promotes coordinated finger closure by structuring the springs stiffness coefficients through a geometric scaling law:

$$k_{ij} = k_0 \, \alpha^{\frac{i-1}{N_\mathrm{f}-1}} \, \beta^{\frac{j-1}{N_\mathrm{p}-1}}, \tag{4}$$

where N_f and N_p are the number of fingers and phalanges, i is the finger index, j is the phalanx index, and k_0 is a base stiffness chosen experimentally. This formulation reduces the stiffness design space to two variables α and β, making it tractable to systematic exploration. The parameter α modulates the relative motion between the fingers, as illustrated in Fig. 3. The parameter β regulates the intra-finger closure pattern by assigning different attraction forces to each phalanx.

The second mechanism ensures that multiple fingers reach the object surface nearly simultaneously, preventing ejection at first contact. This is achieved by introducing distance-dependent damping slowing down each finger during the approach phase. Each spring is paired with a damper whose coefficient follows

$$c\left(|\boldsymbol{z}|\right) = c_0 + c_1 \, \exp\left(-\frac{|\boldsymbol{z}|}{\tau}\right). \tag{5}$$

A uniform damping term c_0 suppresses oscillations and smooths the overall motion, while an exponential term c_1 provides the distance-dependent deceleration near the surface nof the object. The decay constant τ can be re-expressed as $-d/\ln(0.2)$, where d represents the distance at which the exponential damping reaches $20\,\%$ of its maximum value.

As a last step, stable contacts and transient motions must be fine-tuned for the *specific grasps*, namely medium wrap, power sphere, and lateral pinch. All three grasps share the sliding-contact architecture but are further differentiated by grasp-specific elements that encode the particular requirements of each grasp type, as shown in Fig. 2. For the medium wrap, we set α larger than one to achieve thumb–index stabilizing pinch before full-hand closure. For the power sphere, angular springs maintain inter-finger spacing and promote object enclosure. Setting $\alpha = 1$ promotes symmetry in the fingers motion, preventing lateral ejection (see Fig. 3). Finally, for the lateral pinch, both extremities of the index finger are driven toward the box corners, maximizing span. Parameter values for each grasp type are provided in Table I.

TABLE I

PARAMETER VALUES FOR EACH GRASP TYPE.

	k_s	c_max	k_0	α	β	c_0	c_1	d
Medium wrap	5.0	5.0	0.05	1.5	0.5	0.05	0.1	0.01
Power sphere	5.0	5.0	0.05	1.0	0.1	0.15	0.2	0.01
Lateral pinch	5.0	5.0	0.1	1.0	1.0	0.05	0.1	0.005

Fig. 4. The Shadow Dexterous Hand is used to evaluate the robustness of the proposed approach on 27 3D-printed primitive-shaped objects spanning the typical range of daily-life object dimensions.

IV. EXPERIMENTS

A. Experimental setup and practical implementation

Experiments were conducted with a Shadow Dexterous Hand from Shadow Robot [7] mounted on a UR10e arm, and grasping a set of 27 3D-printed primitive shapes covering the typical dimensions of daily life objects. The overall experimental setup is shown in Fig. 4. Throughout all experiments, the objects were placed in favorable dispositions for grasping: cylinders were placed upright, spheres were supported on dedicated stands, and flat objects were placed on elevated surfaces. A grasp is considered successfull if the object is lift up for five seconds without slipping out of the hand.

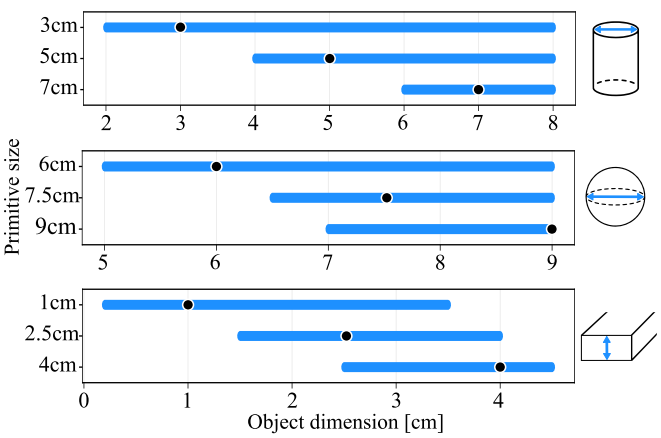

Fig. 5. Dimensional estimation errors of approximately $1\,\mathrm{cm}$ do not compromise grasp success. Blue intervals show the range of successful grasps for each chosen primitive size; black points mark exact matches between primitive and object.

The Shadow Dexterous Hand being a position-controlled robot, direct torque-level VMC control is precluded. The framework is therefore adapted to operate as a grasp planner: a virtual replica of the hand is constructed using virtual mechanisms (links, joints, masses, and inertias), all previously described components are attached to this virtual hand, and the resulting mechanical system is simulated. The joint positions of the virtual hand are then sent as references to the physical robot. All computations[1] are performed using `VMRobotControl.jl` [47], a Julia library for the design, simulation, control, and optimization of VMC architectures.

B. Robustness to object size and pose uncertainty

Robustness to dimension estimation errors is evaluated by choosing three primitive sizes for each grasp type. The associated grasp is applied across a broad range of object dimensions spanning the typical range of human grasping. These three dimensions have been chosen to cover the object test set. Fig. 5 illustrates the existence of a confidence interval around each motion, within which dimensional estimation errors on the order of $1\,\mathrm{cm}$ do not induce any drop in performance. For this experiment, each grasp was repeated five times, consistently yielding identical results. This demonstrates the reliability of the proposed approach.

Robustness against errors in object position and orientation has been evaluated as well. For each grasp type, three meaningful perturbation axes were selected. Perturbations of increasing magnitude were applied across the three previously chosen primitive object dimensions. The results are reported in Fig. 6. They show how the generated grasps can reliably tolerate estimation errors on the order of several centimeters in translation and several tens of degrees in orientation, without significant performance degradation.

Altogether, these two experiments demonstrates the main strength of the proposed method : by integrating uncertainty in simplified primitive models and taking inspiration from human prehension, the resulting motions maintain reliable grasping despite significant errors in object dimensions, position, and orientation.

[1]The code will be released on a GitHub page after the double-blind review process

C. Sensitivity to stiffness and damping parameters

The objective of this experiment is to evaluate the sensitivity of the grasping strategy to parameter tuning. Understanding this sensitivity clarifies how easily the method can be deployed without extensive parameter optimization.

For the nine test objects of Fig. 6, both stiffness and damping parameters were independently perturbed within a $\pm 25\,\%$ range around their nominal values. Variations in the damping parameters did not result in any grasp failures. In contrast, stiffness perturbations produced a single object failure at the boundary of the tested range, corresponding to the largest sphere.

This failure can be traced to the perturbation of the inter-finger stiffness distribution parameter α, whose nominal value is set to 1.0 for the spherical grasp to enforce perfect finger symmetry. Deviating from this symmetric configuration disrupts coordinated enclosure and leads to early object ejection, as illustrated in Fig. 3.

These results highlight how the proposed method does not require refined parameter tuning to achieve reliable grasping performance. Instead, it relies on the relative scaling between mechanical parameters, which can vary within broad ranges without compromising stability or success. This insensitivity to tuning is rooted in the inherent nature of the planner, which leverages the topological coupling of the system's elements rather than their individual numerical values.

D. Baseline comparison

We compare our approach to $\mathcal{D}(\mathcal{R},\mathcal{O})$ grasp [13], a deep learning-based grasp synthesis method that generates grasp poses directly from arbitrary point clouds. Both methods were tested on 3D-printed primitive-shaped objects subject to position perturbations no greater than $0.5\,\mathrm{cm}$. Fig. 7 shows three representative cases. Although both approaches robustly grasp the largest cylinder, the smaller object cases reveal the higher sensitivity of $\mathcal{D}(\mathcal{R},\mathcal{O})$ to position errors. $\mathcal{D}(\mathcal{R},\mathcal{O})$ achieves stable grasp, but the objects deviate significantly from their intended orientation, resulting in less stable final configurations.

This behavior reflects the underlying strategies. $\mathcal{D}(\mathcal{R},\mathcal{O})$ predicts static grasp poses anchored to specific contact points. This precise contact prediction enables complex, object-specific grasps, but reduces tolerance to estimation errors. In contrast, our method generates full, human-inspired grasping trajectories that inherently increase robustness. Taken together, these results suggest that the two approaches address different regions of the grasping problem: our method prioritizes robustness and task-oriented grasps for simple geometries, while $\mathcal{D}(\mathcal{R},\mathcal{O})$ provides object-specific adaptability for complex shapes.

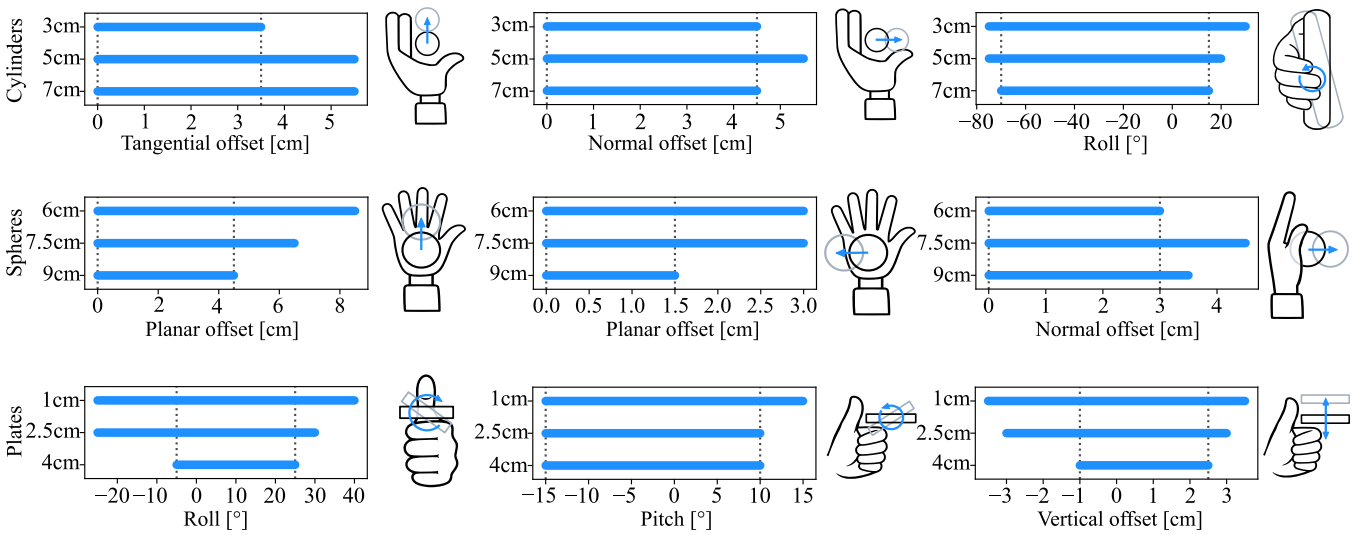

Fig. 6. The method remains reliable despite substantial errors in object pose estimation, tolerating translations of several centimeters and orientations of several tens of degrees across all grasp types and object sizes. Grey dotted lines mark the minimum successful range shared across the three test objects.

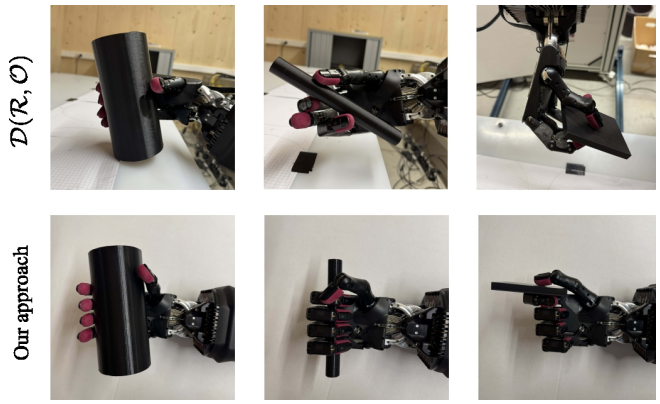

Fig. 7. Adaptive closing trajectories confer robustness that static pose prediction cannot match. For small objects, $\mathcal{D}(\mathcal{R}, \mathcal{O})$ achieves contact but with significant orientation deviation; the proposed method maintains stable, upright grasps under the same perturbations.

E. Real-life objects

As a last step, we assess the ability of our approach to grasp daily-life objects. Despite the coarse object modeling, we argue that the proposed approach can grasp a broader class of objects whose geometry does not significantly deviate from the assumed shape primitives. In practice, the key idea is that objects with approximately isotropic dimensions can be approximated as spheres, objects with one dominant dimension can be modeled as cylinders, and objects with two dominant planar dimensions can be represented as flat boxes.

To evaluate this generalization capability, we tested our approach on a set of 35 objects from the YCB dataset [48], primarily focusing on *Food* and *Kitchen* categories. These were complemented by 45 additional everyday objects, to further extend the test set. All evaluated objects are shown in Fig. 8. The method achieved a success rate of $\mathbf{77}\,\%$ on the YCB subset and $\mathbf{82.5}\,\%$ on the complete object set.

These results show that the proposed approach generalizes beyond the simplified primitive models, successfully grasping objects with more complex geometries. This directly supports the main hypothesis of this paper, that shape primitives enable adaptive grasping without detailed object information, factually demonstrating how several common household objects can be reasonably approximated by three simple primitive categories.

V. Conclusion

This work introduces a grasp planning method specifically designed for robust grasping in unstructured environments. Such robustness is achieved through a template-based strategy, applied both to the object representation (via shape primitive models) and to hand motion (via taxonomy-inspired grasp types). Adaptive behavior within this framework is enabled by Virtual Model Control, which frames the grasp as a mechanical interconnection between the hand and the object. Experimental validation on a Shadow Dexterous Hand demonstrates that the method can handle significant errors in object dimensions, position, and orientation without requiring precise parameter tuning.

The main limitation of this work lies in the pre-grasp phase. Currently, both the object estimation (where primitive size is fixed manually) and the approach phase (where the final hand pose is hard-coded) are pre-defined. These limitations will be addressed in an extended journal version of this work. The goal is to provide a comprehensive grasping pipeline that integrates camera-based object detection with autonomous execution. Future research will also focus on uncertainty compensation through real-time refinement of shape primitives, leveraging compliant control and tactile feedback.

Fig. 8. Shape primitives are expressive enough to cover a broad range of household objects. The method achieves a success rate of 77 % on the YCB subset and 82.5 % on the full set of 80 objects. YCB objects appear on the left; additional everyday objects on the right.

REFERENCES

[1] A. Billard and D. Kragic, "Trends and challenges in robot manipulation," *Science*, vol. 364, no. 6446, p. eaat8414, 2019.

[2] T. Feix, J. Romero, H.-B. Schmiedmayer, A. M. Dollar, and D. Kragic, "The grasp taxonomy of human grasp types," *IEEE Trans. Human-Mach. Syst.*, vol. 46, no. 1, pp. 66–77, 2015.

[3] J. E. Pratt, "Virtual model control of a biped walking robot," Master's thesis, Massachusetts Institute of Technology, Aug. 1995.

[4] J. Pratt, C.-M. Chew, A. Torres, P. Dilworth, and G. Pratt, "Virtual model control: An intuitive approach for bipedal locomotion," *Int. J. Robot. Res.*, vol. 20, no. 2, pp. 129–143, 2001.

[5] D. Larby and F. Forni, "Optimal virtual model control for robotics: Design and tuning of passivity-based controllers," *IEEE Trans. Robot.*, vol. 42, pp. 439–454, 2026.

[6] Y. Zhang, D. Larby, F. Iida, and F. Forni, "Virtual model control for compliant reaching under uncertainties," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2024, pp. 795–801.

[7] A. Kochan, "Shadow delivers first hand," *Ind. Robot*, vol. 32, no. 1, pp. 15–16, 2005.

[8] C. Ferrari, J. Canny *et al.*, "Planning optimal grasps," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, vol. 3. IEEE, 1992, pp. 2290–2295.

[9] A. Bicchi and V. Kumar, "Robotic grasping and contact: A review," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, vol. 1, 2000, pp. 348–353.

[10] J. Bohg, A. Morales, T. Asfour, and D. Kragic, "Data-driven grasp synthesis—A survey," *IEEE Trans. Robot.*, vol. 30, no. 2, pp. 289–309, 2013.

[11] T. G. W. Lum, A. H. Li, P. Culbertson, K. Srinivasan, A. D. Ames, M. Schwager, and J. Bohg, "Get a grip: Multi-finger grasp evaluation at scale enables robust sim-to-real transfer," *arXiv preprint arXiv:2410.23701*, 2024.

[12] J. Varley, C. DeChant, A. Richardson, J. Ruales, and P. Allen, "Shape completion enabled robotic grasping," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2017, pp. 2442–2447.

[13] Z. Wei, Z. Xu, J. Guo, Y. Hou, C. Gao, Z. Cai, J. Luo, and L. Shao, "$\mathcal{D}(\mathcal{R}, \mathcal{O})$ grasp: A unified representation of robot and object interaction for cross-embodiment dexterous grasping," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2025, pp. 4982–4988.

[14] S. Ekvall and D. Kragic, "Learning and evaluation of the approach vector for automatic grasp generation and planning," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2007, pp. 4715–4720.

[15] D. D. Hoffman and W. A. Richards, "Parts of recognition," *Cognition*, vol. 18, no. 1-3, pp. 65–96, 1984.

[16] I. Biederman, "Recognition-by-components: A theory of human image understanding," *Psychol. Rev.*, vol. 94, no. 2, pp. 115–147, 1987.

[17] A. T. Miller, S. Knoop, H. I. Christensen, and P. K. Allen, "Automatic grasp planning using shape primitives," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, vol. 2, 2003, pp. 1824–1829.

[18] K. Huebner, S. Ruthotto, and D. Kragic, "Minimum volume bounding box decomposition for shape approximation in robot grasping," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2008, pp. 1628–1633.

[19] K. Huebner and D. Kragic, "Selection of robot pre-grasps using box-based shape approximation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2008, pp. 1765–1770.

[20] A. Palleschi, F. Angelini, C. Gabellieri, D. W. Park, L. Pallottino, A. Bicchi, and M. Garabini, "Grasp it like a pro 2.0: A data-driven approach exploiting basic shape decomposition and human data for grasping unknown objects," *IEEE Trans. Robot.*, vol. 39, no. 5, pp. 4016–4036, 2023.

[21] C. Goldfeder, P. K. Allen, C. Lackner, and R. Pelossof, "Grasp planning via decomposition trees," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2007, pp. 4679–4684.

[22] G. Vezzani, U. Pattacini, and L. Natale, "A grasping approach based on superquadric models," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2017, pp. 1579–1586.

[23] M. Przybylski, T. Asfour, and R. Dillmann, "Unions of balls for shape approximation in robot grasping," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2010, pp. 1592–1599.

[24] ——, "Planning grasps for robotic hands using a novel object representation based on the medial axis transform," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2011, pp. 1781–1788.

[25] N. Vahrenkamp, E. Koch, M. Waechter, and T. Asfour, "Planning high-quality grasps using mean curvature object skeletons," *IEEE Robot. Autom. Lett.*, vol. 3, no. 2, pp. 911–918, 2018.

[26] Q. Lu and T. Hermans, "Modeling grasp type improves learning-based grasp planning," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 784–791, 2019.

[27] D. Lyons, "A simple set of grasps for a dextrous hand," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, vol. 2, 1985, pp. 588–593.

[28] S. A. Stansfield, "Robotic grasping of unknown objects: A knowledge-based approach," *Int. J. Robot. Res.*, vol. 10, no. 4, pp. 314–326, 1991.

[29] J. Romero, H. Kjellstrom, and D. Kragic, "Modeling and evaluation of human-to-robot mapping of grasps," in *Proc. Int. Conf. Adv. Robot. (ICAR)*, 2009, pp. 1–6.

[30] S. B. Kang and K. Ikeuchi, "Toward automatic robot instruction from perception-mapping human grasps to manipulator grasps," *IEEE Trans. Robot. Autom.*, vol. 13, no. 1, pp. 81–95, 1997.

[31] A. Morales, T. Asfour, P. Azad, S. Knoop, and R. Dillmann, "Integrated grasp planning and visual object localization for a humanoid robot with five-fingered hands," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2006, pp. 5663–5668.

[32] K. Harada, K. Kaneko, and F. Kanehiro, "Fast grasp planning for hand/arm systems based on convex model," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2008, pp. 1162–1168.

[33] Z. Deng, G. Gao, S. Frintrop, F. Sun, C. Zhang, and J. Zhang, "Attention based visual analysis for fast grasp planning with a multi-fingered robotic hand," *Front. Neurorobot.*, vol. 13, p. 60, 2019.

[34] Z. Deng, B. Fang, B. He, and J. Zhang, "An adaptive planning framework for dexterous robotic grasping with grasp type detection," *Robot. Auton. Syst.*, vol. 140, p. 103727, 2021.

[35] H. B. Amor, O. Kroemer, U. Hillenbrand, G. Neumann, and J. Peters, "Generalization of human grasping for multi-fingered robot hands,"

in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2012, pp. 2043–2050.

[36] J. Lundell, E. Corona, T. N. Le, F. Verdoja, P. Weinzaepfel, G. Rogez, F. Moreno-Noguer, and V. Kyrki, "Multi-fingan: Generative coarse-to-fine sampling of multi-finger grasps," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2021, pp. 4495–4501.

[37] D. Dimou, J. Santos-Victor, and P. Moreno, "Grasp pose sampling for precision grasp types with multi-fingered robotic hands," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots (Humanoids)*, 2022, pp. 773–779.

[38] Y. Zhang, J. Hang, T. Zhu, X. Lin, R. Wu, W. Peng, D. Tian, and Y. Sun, "Functionalgrasp: Learning functional grasp for robots via semantic hand-object representation," *IEEE Robot. Autom. Lett.*, vol. 8, no. 5, pp. 3094–3101, 2023.

[39] M. Santello, M. Flanders, and J. F. Soechting, "Postural hand synergies for tool use," *J. Neurosci.*, vol. 18, no. 23, pp. 10 105–10 115, 1998.

[40] M. Santello, M. Bianchi, M. Gabiccini, E. Ricciardi, G. Salvietti, D. Prattichizzo, M. Ernst, A. Moscatelli, H. Jörntell, A. M. Kappers *et al.*, "Hand synergies: Integration of robotics and neuroscience for understanding the control of biological and artificial hands," *Phys. Life Rev.*, vol. 17, pp. 1–23, 2016.

[41] M. Ciocarlie, C. Goldfeder, and P. Allen, "Dexterous grasping via eigengrasps: A low-dimensional approach to a high-complexity problem," in *Proc. Robot.: Sci. Syst. (RSS), Workshop on Sensing and Adapting to the Real World*, 2007.

[42] R. Ortega, A. J. van der Schaft, I. Mareels, and B. Maschke, "Putting energy back in control," *IEEE Control Syst. Mag.*, vol. 21, no. 2, pp. 18–33, 2001.

[43] M. W. Spong, "An historical perspective on the control of robotic manipulators," *Annu. Rev. Control Robot. Auton. Syst.*, vol. 5, pp. 1–31, 2022.

[44] N. Chopra, M. Fujita, R. Ortega, and M. W. Spong, "Passivity-based control of robots: Theory and examples from the literature," *IEEE Control Syst. Mag.*, vol. 42, no. 2, pp. 63–73, Apr. 2022.

[45] C. Secchi, S. Stramigioli, and C. Fantuzzi, *Control of Interactive Robotic Interfaces: A Port-Hamiltonian Approach*, ser. Springer Tracts in Advanced Robotics. Berlin: Springer, 2007, no. volume 29.

[46] I. M. Bullock, T. Feix, and A. M. Dollar, "Finding small, versatile sets of human grasps to span common objects," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2013, pp. 1068–1075.

[47] D. Larby, "VMRobotControl.jl," https://github.com/Cambridge-Control-Lab/VMRobotControl.jl, 2024, accessed: Mar. 2026.

[48] B. Calli, A. Singh, A. Walsman, S. Srinivasa, P. Abbeel, and A. M. Dollar, "The YCB object and model set: Towards common benchmarks for manipulation research," in *Proc. Int. Conf. Adv. Robot. (ICAR)*, 2015, pp. 510–517.