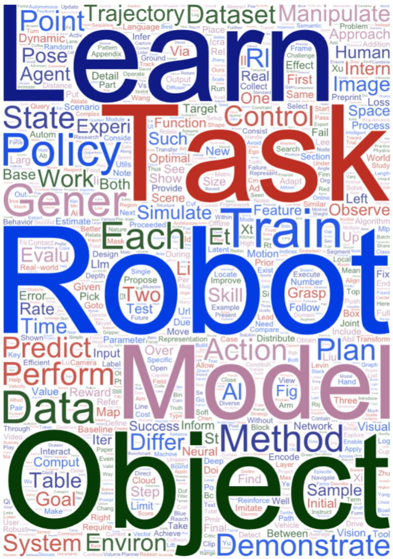



# Robot Learning

## Manipulation & Grasp Learning

Marc Toussaint
Technical University of Berlin
Summer 2024

## Outline

- Manipulation Intro
- Background on Grasping
- Grasp Learning Methods
- Briefly: Other Manipulation Learning

# Manipulation is a Core Challenge in Robotics!

# Manipulation is a Core Challenge in Robotics!

- Recall the “Robotics Essentials Lecture”
  - Robotics is about Articulated Multibody Systems
  - Objects in the environment are part of the “multibody system” (slide 21); have their own DOFs, but are not articulated
  - hybrid dynamics: on-off switching of manipulability; friction, stiction, slip, non-point contacts

# Manipulation is a Core Challenge in Robotics!

- Recall the "Robotics Essentials Lecture"
  - Robotics is about Articulated Multibody Systems
  - Objects in the environment are part of the "multibody system" (slide 21); have their own DOFs, but are not articulated
  - hybrid dynamics: on-off switching of manipulability; friction, stiction, slip, non-point contacts

- Think back about the last 5 lectures & exercises
  - dynamics learning, imitation learning, RL, InvRL, safe learning
  - Most work: state space $\leftrightarrow$ robot configuration (Hopper, Walker, helicopter, UAVs, quadropeds)
  - Few works involved game environments: SpaceInvaders, Pong
  - Some works about image-based manipulation of single object: image $\leftrightarrow$ state

# Manipulation – Definition

- Matt Mason:

  *Manipulation is when an agent moves things other than itself.*

  M. T. Mason. Toward Robotic Manipulation.
  *Annual Review of Control, Robotics, and Autonomous Systems*, 1(1):1–28, 2018-05-28.
  URL: https://www.annualreviews.org/delete_doi/10.1146/annurev-control-060117-104848

# Manipulation – Definition

- Matt Mason:

  *Manipulation is when an agent moves things other than itself.*

  M. T. Mason. Toward Robotic Manipulation.
  *Annual Review of Control, Robotics, and Autonomous Systems*, 1(1):1–28, 2018-05-28.
  URL: https://www.annualreviews.org/delete_doi/10.1146/annurev-control-060117-104848

- My view: *General-purpose Manipulation* $\leftrightarrow$ *Ability to reach any physically possible environment configuration*

# Manipulation – Definition

- Matt Mason:

  *Manipulation is when an agent moves things other than itself.*

  M. T. Mason. Toward Robotic Manipulation.
  *Annual Review of Control, Robotics, and Autonomous Systems*, 1(1):1–28, 2018-05-28.
  URL: `https://www.annualreviews.org/delete_doi/10.1146/annurev-control-060117-104848`

- My view: *General-purpose Manipulation* $\leftrightarrow$ *Ability to reach any physically possible environment configuration*

- Earlier work/definitions was fully focussed on grasping; now includes pushing, throwing, sticking, tools, ropes, any means...
- Great Lecture:

  R. Tedrake. Robotic Manipulation - Lecture Website, 2023.
  URL: `https://manipulation.csail.mit.edu/index.html`

# Manipulation Learning

- What is learned?

**environment/task parameters**

instructions/lang./goal info $g$
physics parameters $\Theta$

**state evaluations**

rewards $r_t$
value $V(x)$
Q-value $Q(x,u)$
constraint $\phi(x)$

state
$x_t$

controls
$u_t$

observations
$y_t$

**plans/anticipation**

waypoints/subgoals $x_{t_{1:K}}$
trajectory $x_{[t,t+H]}$
action plan $a_{1:K}$

# Manipulation Learning

- What is learned?
- Policy: Image $\to$ Controls
  - Grounded in MDP formalism: $x_t, u_t \mapsto r_t, x_{t+1}$
  - is about the control process in fine time resolution

**environment/task parameters**

instructions/lang./goal info $g$
physics parameters $\Theta$

**state evaluations**

rewards $r_t$
value $V(x)$
Q-value $Q(x,u)$
constraint $\phi(x)$

state $x_t$

controls $u_t$

observations $y_t$

**plans/anticipation**

waypoints/subgoals $x_{t_{1:K}}$
trajectory $x_{[t,t+H]}$
action plan $a_{1:K}$

environment/task parameters

instructions/lang./goal info $g$
physics parameters $\Theta$

# Manipulation Learning

state evaluations

rewards $r_t$
value $V(x)$
Q-value $Q(x,u)$
constraint $\phi(x)$

state $x_t$

controls $u_t$

observations $y_t$

plans/anticipation

waypoints/subgoals $x_{t_{1:K}}$
trajectory $x_{[t,t+H]}$
action plan $a_{1:K}$

- What is learned?

- Policy: Image $\to$ Controls
  - Grounded in MDP formalism: $x_t, u_t \mapsto r_t, x_{t+1}$
  - is about the control process in fine time resolution

- Solutions/Constraints: Image $\to$ grasp pose, push pose
  - Not about the control process; no MDP formalism; no rewards, but $x \mapsto$success/no-success
  - The learned model predicts successful grasps, push poses, throw parameters, etc
  - These are then executed using standard control theory

# Outline

- Manipulation Intro
- **Background on Grasping**
- Grasp Learning Methods
- Briefly: Other Manipulation Learning

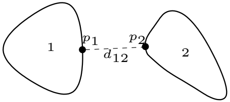
# Grasping Background

See also Chapter 12 of

K. M. Lynch and F. C. Park. *Modern Robotics*.
2017.
URL: https://books.google.com/books?hl=en&lr=&id=5NzFDgAAQBAJ&oi=fnd&pg=PR11&dq=modern+robotics+book&ots=qsJmY4kXPh&sig=o1uhr6h_eJKF33_HBe2xZaT32Ow

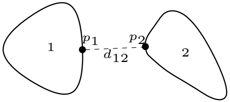
# Contacts

- Contact between two bodies – definitions:
  - configuration $q = (q_1, q_2)$ (with $q_i \in \mathrm{SE}(3)$ pose of $i$th body)
  - Their shapes define the **pairwise signed-distance** $d_{12}(q_1, q_2)$ (and its gradient)
  - Two nearest points $p_1$, $p_2$ are called **witness points**
  - We also have the contact normal $n \in \mathbb{R}^3$



# Contacts

- Contact between two bodies – definitions:
  - configuration $q = (q_1, q_2)$ (with $q_i \in \mathrm{SE}(3)$ pose of $i$th body)
  - Their shapes define the **pairwise signed-distance** $d_{12}(q_1, q_2)$ (and its gradient)
  - Two nearest points $p_1$, $p_2$ are called **witness points**
  - We also have the contact normal $n \in \mathbb{R}^3$




- Multiple contact forces on one body:
  - One body, $C$ contact points at position $p_i$, each creates **wrench** $(f_i, \tau_i) \in \mathbb{R}^6$ at $p_i$, totals:

$$f^{\text{total}} = \sum_{i=1}^{C} f_i \;, \quad \tau^{\text{total}} = \sum_{i=1}^{C} \tau_i + f_i \times (p_i - c)$$

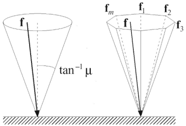
  - Newton-Euler equation describes the resulting acceleration:

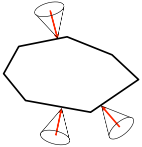
$$\begin{pmatrix} f^{\text{total}} \\ \tau^{\text{total}} \end{pmatrix} = \begin{pmatrix} m\dot{v} \\ I\dot{w} + w \times Iw \end{pmatrix}$$

Since *"Manipulation is when an agent moves things other than itself"* these equations "fully describe" what manipulation is about: Creating contact forces to appropriately accelerate objects.

# Contacts

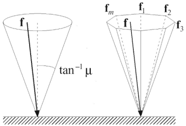
- Contact Friction:
  - Point finger can not transmit torque $\Rightarrow \quad \tau_i = 0$ (better: patch models)
  - Point finger sticks only when tangentil force $f^= \le \mu f^\perp \quad (f^\perp = nn^\top f,\ f^= = f - f^\perp)$
  - The set $F_i = \{f_i : f_i^= \le \mu f_i^\perp\}$ is called the **friction cone**

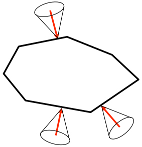


# Contacts

- Contact Friction:
  - Point finger can not transmit torque $\Rightarrow \quad \tau_i = 0$ (better: patch models)
  - Point finger sticks only when tangentil force $f^= \le \mu f^\perp \quad (f^\perp = nn^\top f,\ f^= = f - f^\perp)$
  - The set $F_i = \{f_i : f_i^= \le \mu f_i^\perp\}$ is called the **friction cone**



- **Force closure:**
  - A **contact configuration** $\{(p_i, n_i)\}_{i=1}^C$ with friction coeff $\mu$ creates force closure
    - $\Leftrightarrow$ we can generate (counter-act) arbitrary $f^{\text{total}}$ and $\tau^{\text{total}}$ by choosing $f_i \in F_i$ appropriately.
    - $\Leftrightarrow$ The *positive linear span of the fiction cones* covers the whole space of $(f^{\text{total}}, \tau^{\text{total}}) \in \mathbb{R}^6$

# Force Closure & Force Closure Metric & Form Closure & Caging

- Force closure: The contacts can apply an arbitrary wrench (=force-torque) to the object.

# Force Closure & Force Closure Metric & Form Closure & Caging

- Force closure: The contacts can apply an arbitrary wrench (=force-torque) to the object.
- Force closure metric: Limit finger force $|f_i| \leq 1$ and compute radius (=origin-distance) of convex hull
- Form closure: The object is at an isolated point in configuration space. Note: form closure $\Leftrightarrow$ frictionless force closure
- Caging: The object is not fixated, but cannot escape

## Outline

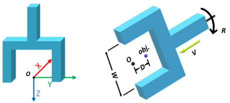
- Manipulation Intro
- Background on Grasping
- **Grasp Learning Methods**
- Briefly: Other Manipulation Learning

# Grasp Learning

- What is learned?

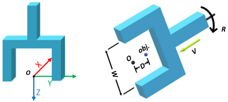
# Grasp Learning

- What is learned?
  - Simplified parallel gripper:
  - Input: RGB-D image of scene

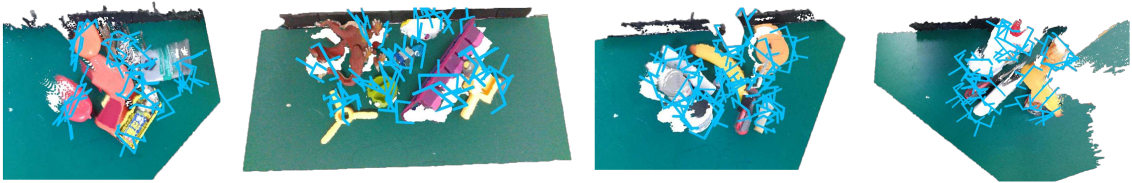


# Grasp Learning

- What is learned?
  - Simplified parallel gripper:



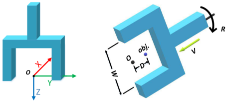

  - Input: RGB-D image of scene
  - Output: Set of **grasps** (=gripper poses $q^{\text{gripper}} \in \text{SE}(3)$) in the scene:



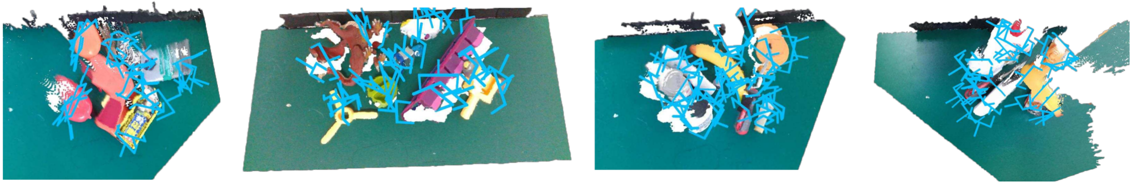
  - Alternative output: A network that can score any proposed grasp

# Grasp Learning

- What is learned?
  - Simplified parallel gripper:




  - Input: RGB-D image of scene
  - Output: Set of **grasps** (=gripper poses $q^{\text{gripper}} \in \text{SE}(3)$) in the scene:



  - Alternative output: A network that can score any proposed grasp

- Training data: pairs of scene (usually converted to **point cloud** $P_s$) and grasps

$$D = \{ \ (P_s, \{q_{s,i}\}_{i=1}^{G_s}) \ \}_{s=1}^{S}$$

# GraspNet 1

**GraspNet-1Billion: A Large-Scale Benchmark for General Object Grasping**

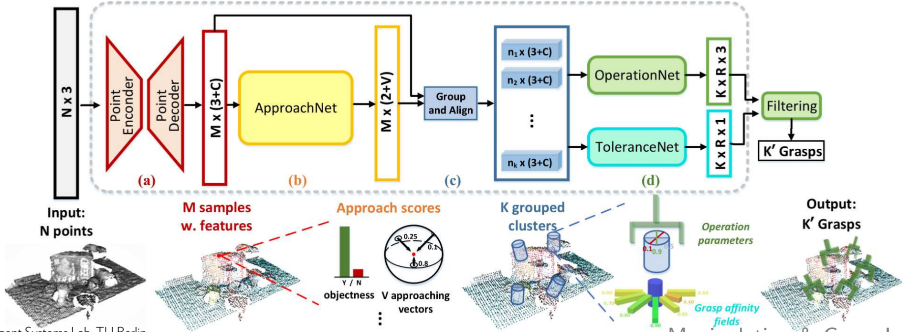
Hao-Shu Fang, Chenxi Wang, Minghao Gou. Cewu Lu[1]
Shanghai Jiao Tong University
fhaoshu@gmail.com, {wcx1997,gmh2015,lucewu}@sjtu.edu.cn

H.-S. Fang, C. Wang, M. Gou, and C. Lu. Graspnet-1billion: A large-scale benchmark for general object grasping.
In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11444–11453, 2020.
URL: http://openaccess.thecvf.com/content_CVPR_2020/html/Fang_GraspNet-1Billion_A_Large-Scale_Benchmark_for_General_Object_Grasping_CVPR_2020_paper.html

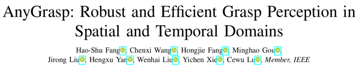
- Focusses on data collection (details later)

$$D = \{(P, \{(\underbrace{p \in P, v, D, R}_{q^{\text{gripper}} \in \text{SE}(3)}, w)_i\})\}$$

- Given data, they propose architecture
  - First PCL $\to v$/success classifier per point $p$
  - Then predict $D, R, w$
  - with separate loss functions for each part



# GraspNet 2


AnyGrasp: Robust and Efficient Grasp Perception in
Spatial and Temporal Domains

Hao-Shu Fang, Chenxi Wang, Hongjie Fang, Minghao Gou,
Jirong Liu, Hengxu Yan, Wenhai Liu, Yichen Xie, Cewu Lu, *Member, IEEE*

H.-S. Fang, C. Wang, H. Fang, M. Gou, J. Liu, H. Yan, W. Liu, Y. Xie, and C. Lu.

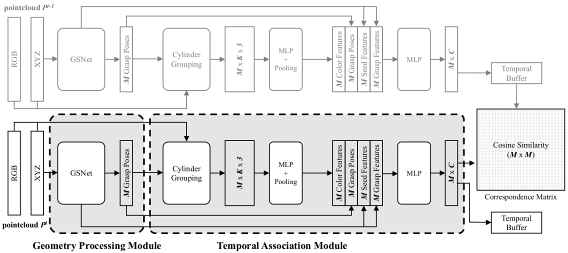
Anygrasp: Robust and efficient grasp perception in spatial and temporal domains.
*IEEE Transactions on Robotics*, 2023.
URL: https://ieeexplore.ieee.org/abstract/document/10167687/

- Much more complex architecture

  https://youtu.be/dNnLgAGreec

- Also dynamic (temporally stable) predictions:

  https://www.youtube.com/watch?v=207Uo0xeLlk



# Other Grasp Learning Work

- Classic: Identifying “antipodal” grasps in point clouds:

  A. Ten Pas, M. Gualtieri, K. Saenko, and R. Platt. Grasp Pose Detection in Point Clouds. *The International Journal of Robotics Research*, 36(13-14):1455–1473, 2017-12.
  URL: http://journals.sagepub.com/delete_doi/10.1177/0278364917735594

- Classic: DexNet family:

  J. Mahler, J. Liang, S. Niyaz, M. Laskey, R. Doan, X. Liu, J. A. Ojea, and K. Goldberg. Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics, 2017-08-08.
  URL: http://arxiv.org/abs/1703.09312, arXiv:1703.09312
  https://www.youtube.com/watch?v=i6K3GI2_EgU

- More from the “RL” side (“closed loop grasping”):

  S. Song, A. Zeng, J. Lee, and T. Funkhouser. Grasping in the wild: Learning 6dof closed-loop grasping from low-cost demonstrations. *IEEE Robotics and Automation Letters*, 5(3):4978–4985, 2020.
  URL: https://ieeexplore.ieee.org/abstract/document/9126187/
  https://www.youtube.com/watch?v=UPJjpIhXpZ8

- Contact-GraspNet

  M. Sundermeyer, A. Mousavian, R. Triebel, and D. Fox. Contact-graspnet: Efficient 6-dof grasp generation in cluttered scenes. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 13438–13444, 2021.
  URL: https://ieeexplore.ieee.org/abstract/document/9561877/
  https://www.youtube.com/watch?v=qRLKYSLXElM

- Using Diffusion Models

  J. Urain, N. Funk, J. Peters, and G. Chalvatzaki. Se (3)-diffusionfields: Learning smooth cost functions for joint grasp and motion optimization through diffusion. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5923–5930, 2023.
  URL: https://ieeexplore.ieee.org/abstract/document/10161569/
  https://www.youtube.com/watch?v=Tk6l3WsPGMY

# Grasp Data Collection

- My view:
  - All of the above papers show: If we have good data, we have good ideas on how to design ML architectures to predict grasps
  - Data Collection is the key!

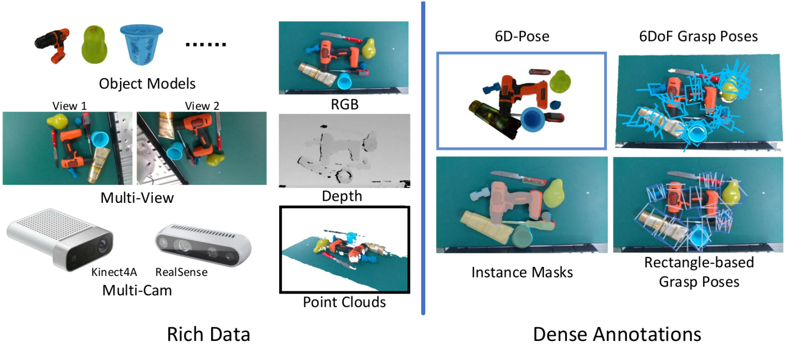
## Grasp Data Collection

- My view:
  - All of the above papers show: If we have good data, we have good ideas on how to design ML architectures to predict grasps
  - Data Collection is the key!

- Two approaches:
  - Model-based labels (grasp theory, force closure)
  - Simulation-based labels

# Model-based Grasp Labels

- GraspNet-1Billion and DexNet 2.0 papers:
  - For every point in the scene, for every (or sampled) approach direction, every offset/roll/width
  - Compute a classical grasp score: Force closure metric
  - Requires knowledge of ground truth object poses and shapes $\rightarrow$ precise object pose estimation

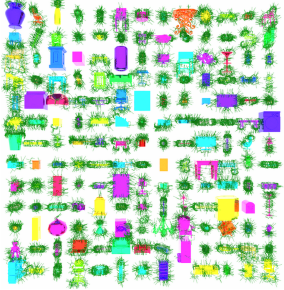


## Model-based Grasp Labels

- So, force closure theory is the origin of wisdom here!
- The learning machinery “only” transfers it to the real world – predicting force closure grasps based on real RGB-D
- Cp. to imitation learning from a privileged expert! Here the privileged expert is the force closure metric assuming known object shapes.

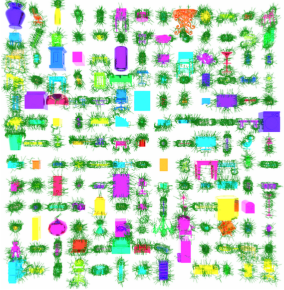
# Simulation-based Grasp Labels

C. Eppner, A. Mousavian, and D. Fox. Acronym: A large-scale grasp dataset based on simulation. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6222–6227, 2021. URL: `https://ieeexplore.ieee.org/abstract/document/9560844/`



- Use generic rigid body physics simulator:
  - Throw random objects (from `ShapeNet`) into a scene (and render RGB-D image)
  - generate random grasps – smartly engineered!
  - Close and lift gripper – measure in-hand motion during both phases
  - "we simulate 17.744 million grasps, out of which 59.21% (ap- proximately 10.5 million grasps) succeed."

# Simulation-based Grasp Labels



C. Eppner, A. Mousavian, and D. Fox. Acronym: A large-scale grasp dataset based on simulation. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6222–6227, 2021. URL: `https://ieeexplore.ieee.org/abstract/document/9560844/`

- Use generic rigid body physics simulator:
  - Throw random objects (from `ShapeNet`) into a scene (and render RGB-D image)
  - generate random grasps – smartly engineered!
  - Close and lift gripper – measure in-hand motion during both phases
  - "we simulate 17.744 million grasps, out of which 59.21% (ap- proximately 10.5 million grasps) succeed."
- So, the physics simulator (=Newton-Euler equations + contact models) is the origin of wisdom here!
  - Again, cp. to imitation learning from privileged expert (=simulation)

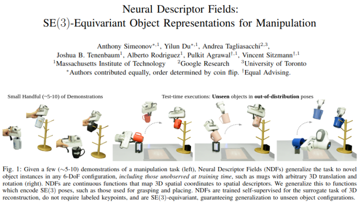
# Grasp Learning Summary

- Rather advanced for standard parallel gripper; less for more complex hands
- In my view, proper data generation is key – existing methods still have deficits
- Given proper data, the advances in learning are unstoppable (stronger architectures, diffusion, etc)

## Manipulation Learning

- Manipulation is more than “pick-and-place”
  - manipulating articulated objects
  - pushing, throwing
  - rolling, spinning, balancing/stacking, etc.

# Recall: Extracting Constraints in Imitation Learning

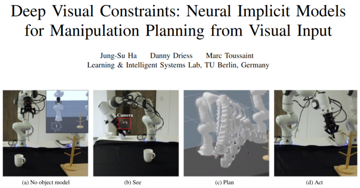


Neural Descriptor Fields:
SE(3)-Equivariant Object Representations for Manipulation

Anthony Simeonov[*,1], Yilun Du[*,1], Andrea Tagliasacchi[2,3],
Joshua B. Tenenbaum[1], Alberto Rodriguez[1], Pulkit Agrawal[†,1], Vincent Sitzmann[†,1]
[1]Massachusetts Institute of Technology [2]Google Research [3]University of Toronto
[*]Authors contributed equally, order determined by coin flip. [†]Equal Advising.



Fig. 1: Given a few (~5-10) demonstrations of a manipulation task (left), Neural Descriptor Fields (NDFs) generalize the task to novel object instances in any 6-DoF configuration, *including those unobserved at training time*, such as mugs with arbitrary 3D translation and rotation (right). NDFs are continuous functions that map 3D spatial coordinates to spatial descriptors. We generalize this to functions which encode SE(3) poses, such as those used for grasping and placing. NDFs are trained self-supervised for the surrogate task of 3D reconstruction, do not require labeled keypoints, and are SE(3)-equivariant, guaranteeing generalization to unseen object configurations.

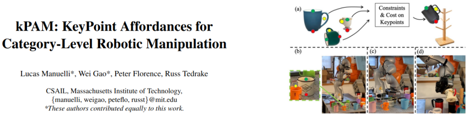


Deep Visual Constraints: Neural Implicit Models for Manipulation Planning from Visual Input

Jung-Su Ha Danny Driess Marc Toussaint
Learning & Intelligent Systems Lab, TU Berlin, Germany



(a) No object model (b) See (c) Plan (d) Act



**kPAM: KeyPoint Affordances for Category-Level Robotic Manipulation**

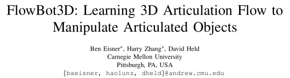
Lucas Manuelli*, Wei Gao*, Peter Florence, Russ Tedrake

CSAIL, Massachusetts Institute of Technology,
{manuelli, weigao, peteflo, russt}@mit.edu
*These authors contributed equally to this work.*



- Extract "constraints of success", but eventually pick-and-place

# Manipulating Learning for Articulated Objects


FlowBot3D: Learning 3D Articulation Flow to
Manipulate Articulated Objects

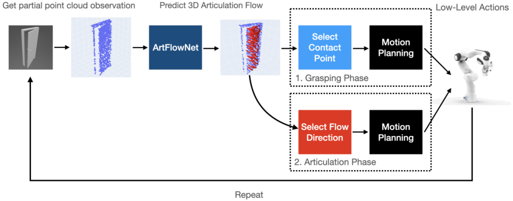
Ben Eisner*, Harry Zhang*, David Held
Carnegie Mellon University
Pittsburgh, PA, USA
{baeisner, haolunz, dheld}@andrew.cmu.edu

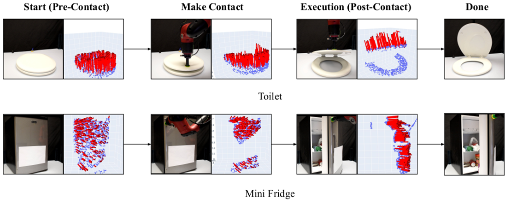
B. Eisner, H. Zhang, and D. Held. FlowBot3D: Learning 3D Articulation Flow to Manipulate Articulated Objects, 2024-05-02.
URL: `http://arxiv.org/abs/2205.04382`, `arXiv:2205.04382`

- Assumes "gripper can be attached to any point on surface"
- Learn a mapping $P \mapsto$ flow field $F_p \in \mathbb{R}^3$ for each $p \in P$

`https://drive.google.com/file/d/`

`1jiEHT--WQec5diEJE6a4dMJkBnP3d36B/view`



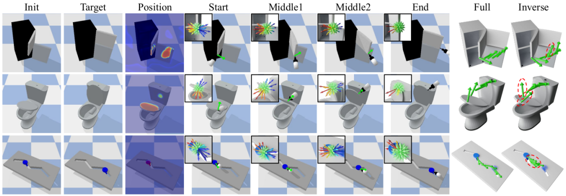




- Similar earlier work:

**UMPNet: Universal Manipulation Policy Network for Articulated Objects**

Zhenjia Xu   Zhanpeng He   Shuran Song
Columbia University
https://ump-net.cs.columbia.edu/

Z. Xu, Z. He, and S. Song. Universal manipulation policy network for articulated objects.
*IEEE robotics and automation letters*, 7(2):2447–2454, 2022.
URL: `https://ieeexplore.ieee.org/abstract/document/9681198/`



# Conclusions

- Manipulation Learning is often beyond the MDP and RL framework!
- We often don't learn low-level policies, but:
  - Predicting grasps in an RGB-D scene
  - Predicting manipulability (flow) of articulated objects from RGB-D
  - Predicting keypoints/waypoints of interaction

# Conclusions

- Manipulation Learning is often beyond the MDP and RL framework!
- We often don't learn low-level policies, but:
  - Predicting grasps in an RGB-D scene
  - Predicting manipulability (flow) of articulated objects from RGB-D
  - Predicting keypoints/waypoints of interaction

- BUT, I think this is sooo far away from truely understanding/learning General-purpose Manipulation!

[1] B. Eisner, H. Zhang, and D. Held.
FlowBot3D: Learning 3D Articulation Flow to Manipulate Articulated Objects, 2024-05-02.
URL: http://arxiv.org/abs/2205.04382, arXiv:2205.04382.

[2] C. Eppner, A. Mousavian, and D. Fox.
Acronym: A large-scale grasp dataset based on simulation.
In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6222–6227, 2021.
URL: https://ieeexplore.ieee.org/abstract/document/9560844/.

[3] H.-S. Fang, C. Wang, H. Fang, M. Gou, J. Liu, H. Yan, W. Liu, Y. Xie, and C. Lu.
Anygrasp: Robust and efficient grasp perception in spatial and temporal domains.
*IEEE Transactions on Robotics*, 2023.
URL: https://ieeexplore.ieee.org/abstract/document/10167687/.

[4] H.-S. Fang, C. Wang, M. Gou, and C. Lu.
Graspnet-1billion: A large-scale benchmark for general object grasping.
In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11444–11453, 2020.
URL: http://openaccess.thecvf.com/content_CVPR_2020/html/Fang_GraspNet-1Billion_A_Large-Scale_Benchmark_for_General_Object_Grasping_CVPR_2020_paper.html.

[5] K. M. Lynch and F. C. Park.
*Modern Robotics*.
2017.
URL: https://books.google.com/books?hl=en&lr=&id=5NzFDgAAQBAJ&oi=fnd&pg=PR11&dq=modern+robotics+book&ots=qsJmY4kXPh&sig=o1uhr6h_eJKF33_HBe2xZaT32Ow.

[6] J. Mahler, J. Liang, S. Niyaz, M. Laskey, R. Doan, X. Liu, J. A. Ojea, and K. Goldberg.
Dex-Net 2.0: Deep Learning to Plan Robust Grasps with Synthetic Point Clouds and Analytic Grasp Metrics, 2017-08-08.
URL: http://arxiv.org/abs/1703.09312, arXiv:1703.09312.

[7] M. T. Mason.
Toward Robotic Manipulation.
*Annual Review of Control, Robotics, and Autonomous Systems*, 1(1):1–28, 2018-05-28.
URL: https://www.annualreviews.org/delete_doi/10.1146/annurev-control-060117-104848.

[8] S. Song, A. Zeng, J. Lee, and T. Funkhouser.
Grasping in the wild: Learning 6dof closed-loop grasping from low-cost demonstrations.
*IEEE Robotics and Automation Letters*, 5(3):4978–4985, 2020.
URL: https://ieeexplore.ieee.org/abstract/document/9126187/.

[9] M. Sundermeyer, A. Mousavian, R. Triebel, and D. Fox.
Contact-graspnet: Efficient 6-dof grasp generation in cluttered scenes.
In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 13438–13444, 2021.
URL: https://ieeexplore.ieee.org/abstract/document/9561877/.

[10] R. Tedrake.
Robotic Manipulation - Lecture Website, 2023.
URL: https://manipulation.csail.mit.edu/index.html.

[11] A. Ten Pas, M. Gualtieri, K. Saenko, and R. Platt.
Grasp Pose Detection in Point Clouds.
*The International Journal of Robotics Research*, 36(13-14):1455–1473, 2017-12.
URL: http://journals.sagepub.com/delete_doi/10.1177/0278364917735594.

[12] J. Urain, N. Funk, J. Peters, and G. Chalvatzaki.
Se (3)-diffusionfields: Learning smooth cost functions for joint grasp and motion optimization through diffusion.
In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5923–5930, 2023.
URL: https://ieeexplore.ieee.org/abstract/document/10161569/.

[13] Z. Xu, Z. He, and S. Song.

Universal manipulation policy network for articulated objects.
*IEEE robotics and automation letters*, 7(2):2447–2454, 2022.
URL: `https://ieeexplore.ieee.org/abstract/document/9681198/`.