Shape2Pose:
Human-Centric Shape Analysis

Vladimir G. Kim[1]

Siddhartha Chaudhuri[2]

Leonidas Guibas[1]

Thomas Funkhouser[2]

1

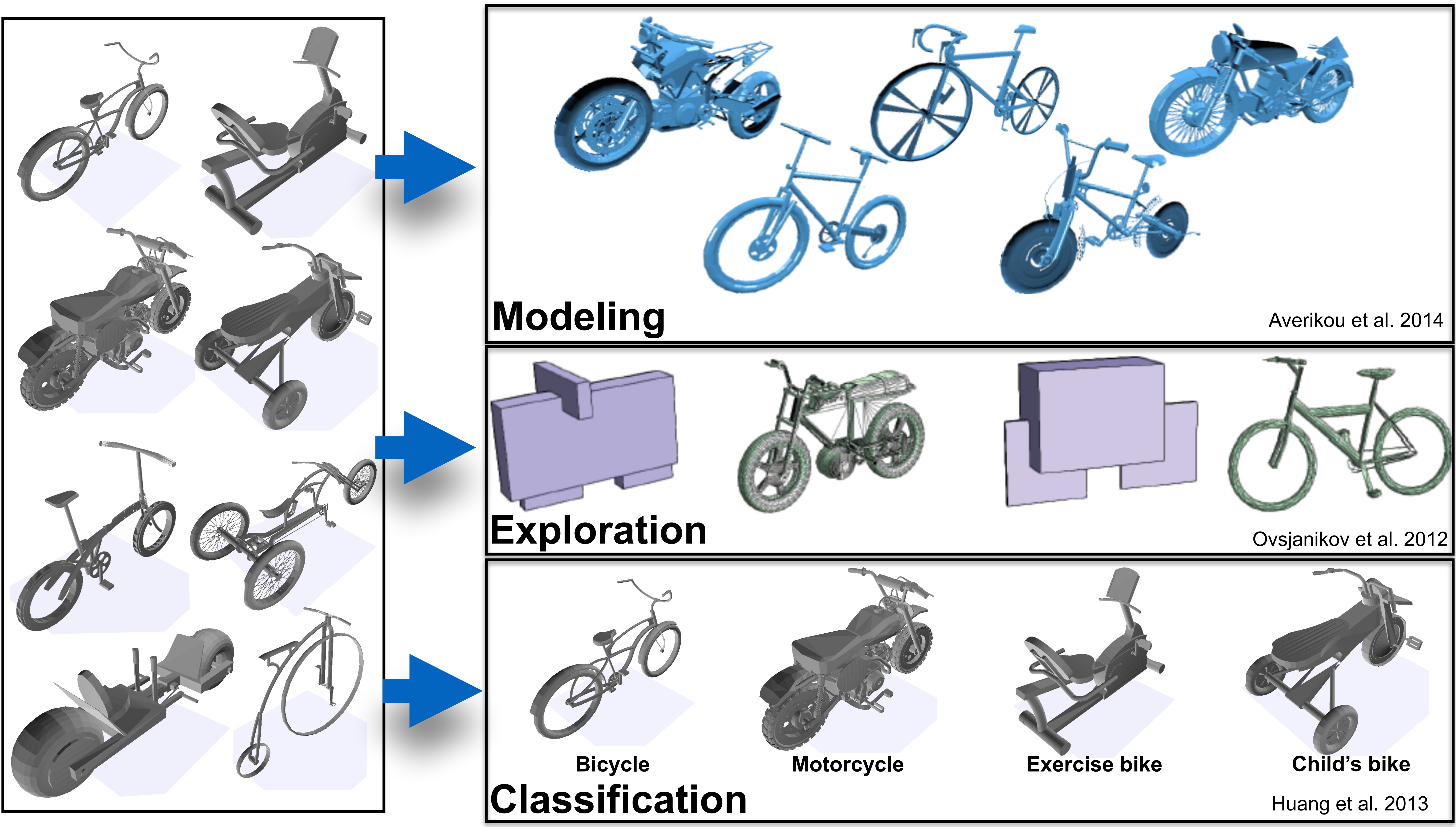
Stanford University

2

Princeton University

Goal

Leverage online repositories to understand classes of shapes



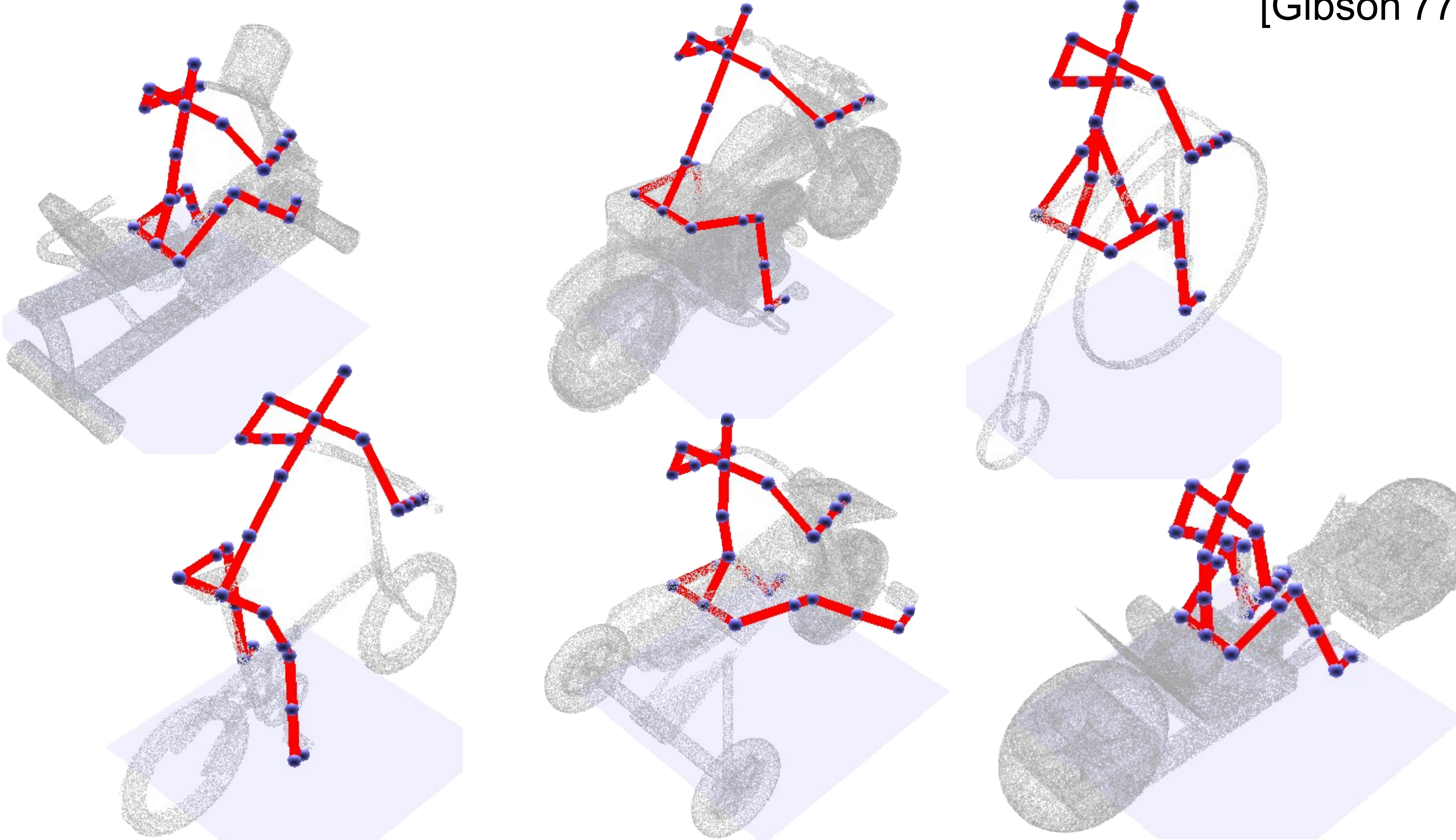
Challenge

Find common structure

Key Idea

Affordance is an intrinsic property of a shape

[Gibson'77]

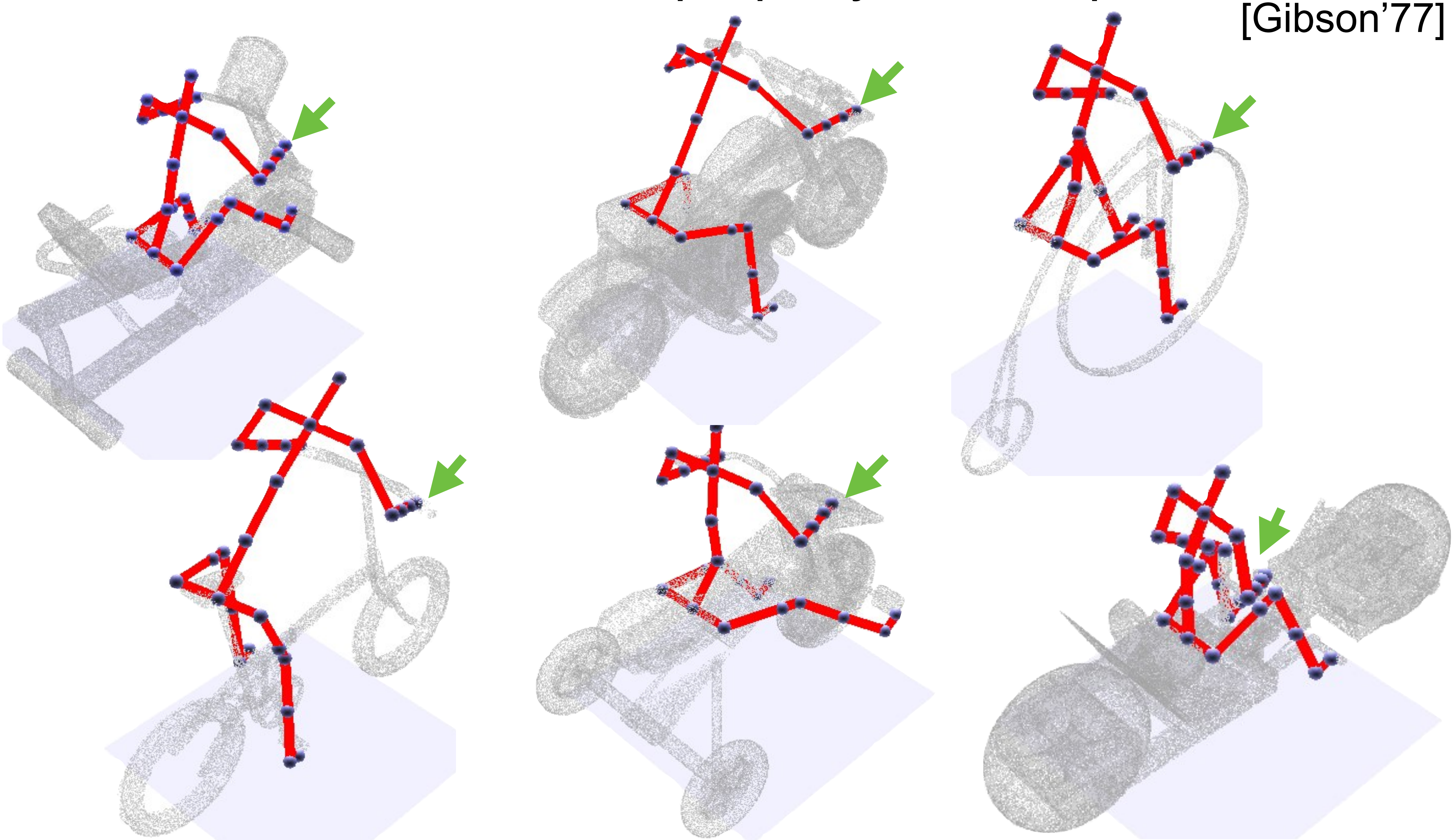


Key Idea

✓ Point-to-point Corrs
✷ Functional Parts
✷ Structural Variations

Affordance is an intrinsic property of a shape

[Gibson'77]

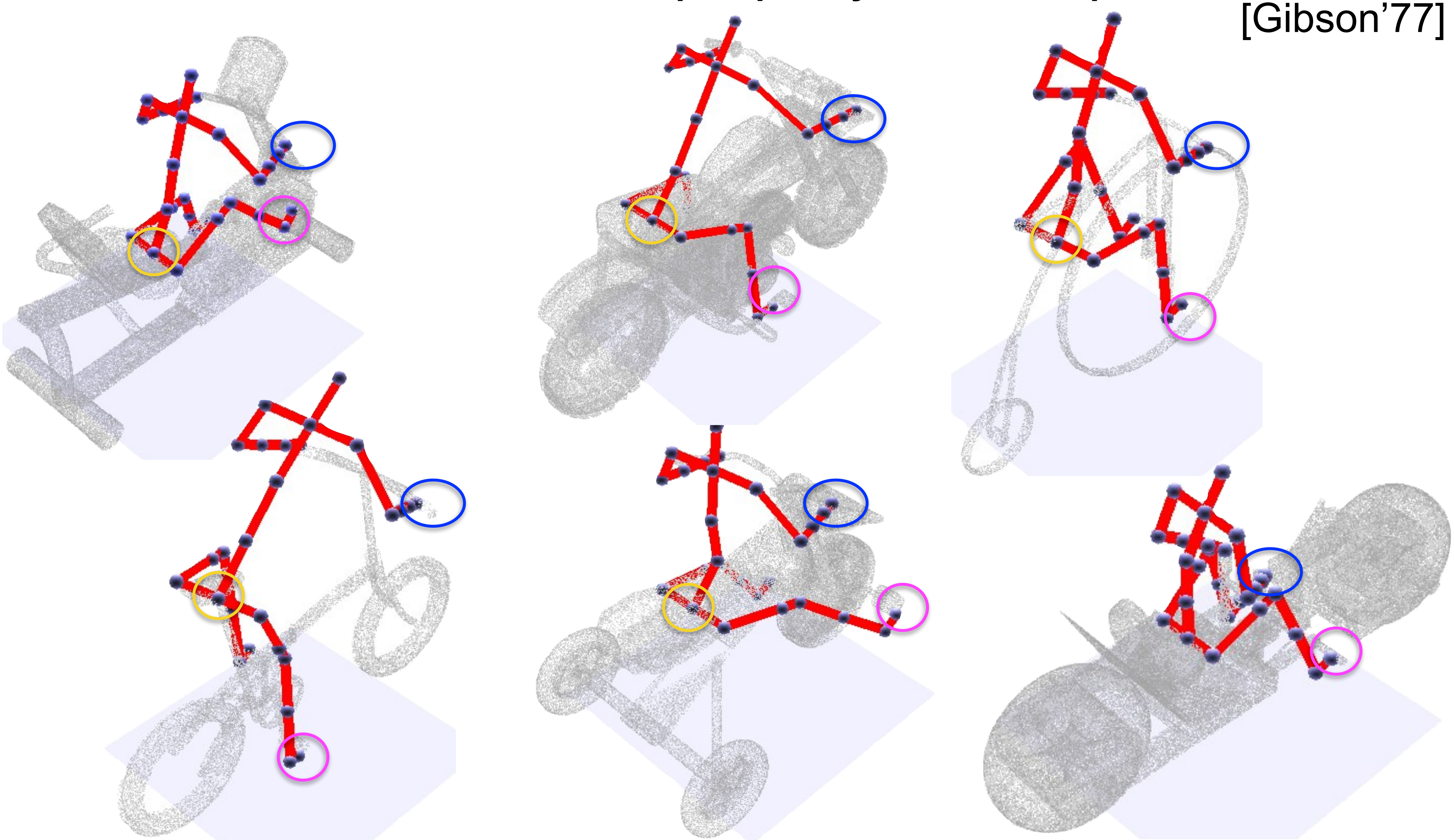


Key Idea

✓ Point-to-point Corrs
✓ Functional Parts
✹ Structural Variations

Affordance is an intrinsic property of a shape

[Gibson'77]

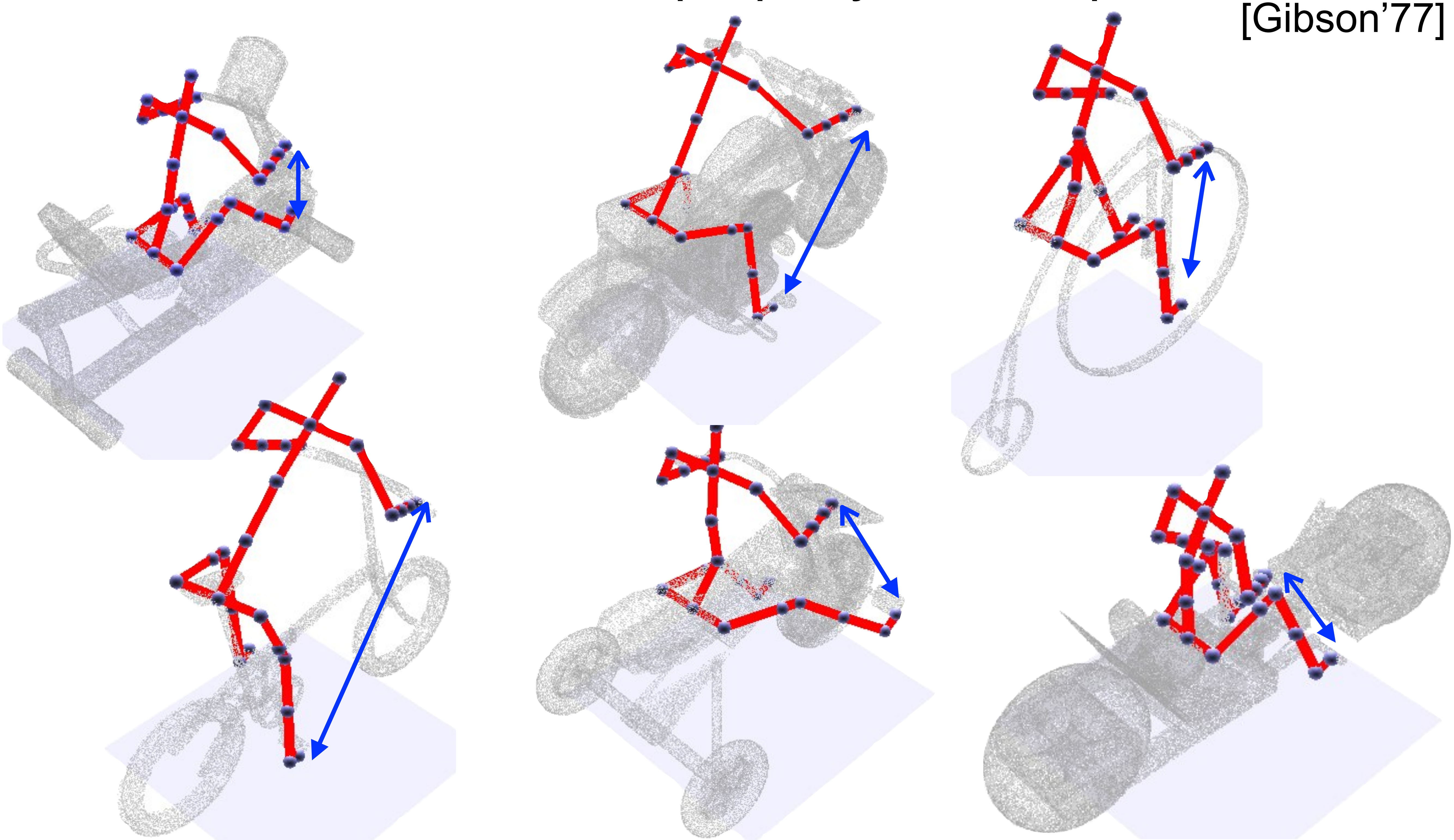


Key Idea

✓ Point-to-point Corrs
✓ Functional Parts
✓ Structural Variations

Affordance is an intrinsic property of a shape

[Gibson'77]

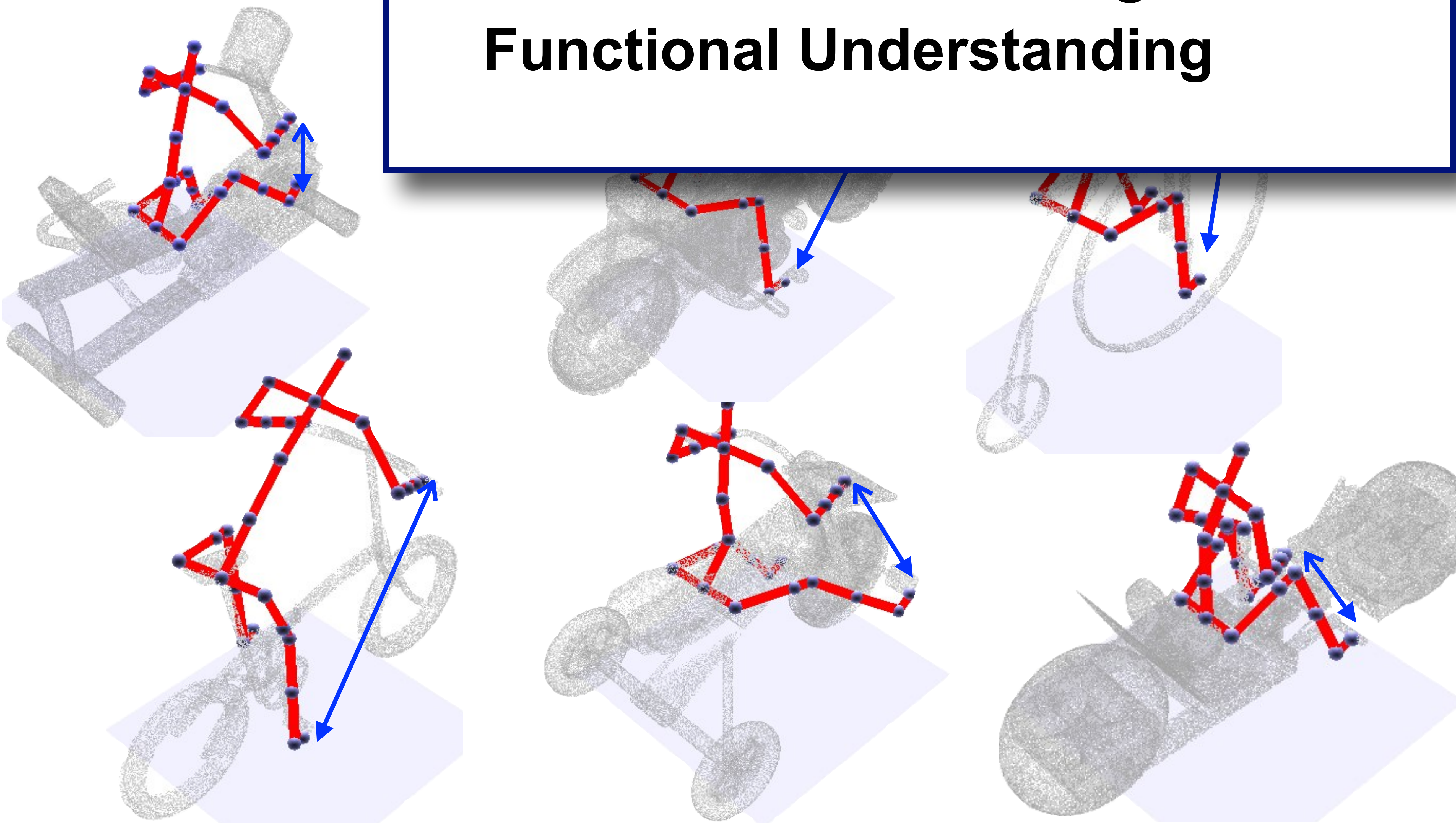


Key Idea

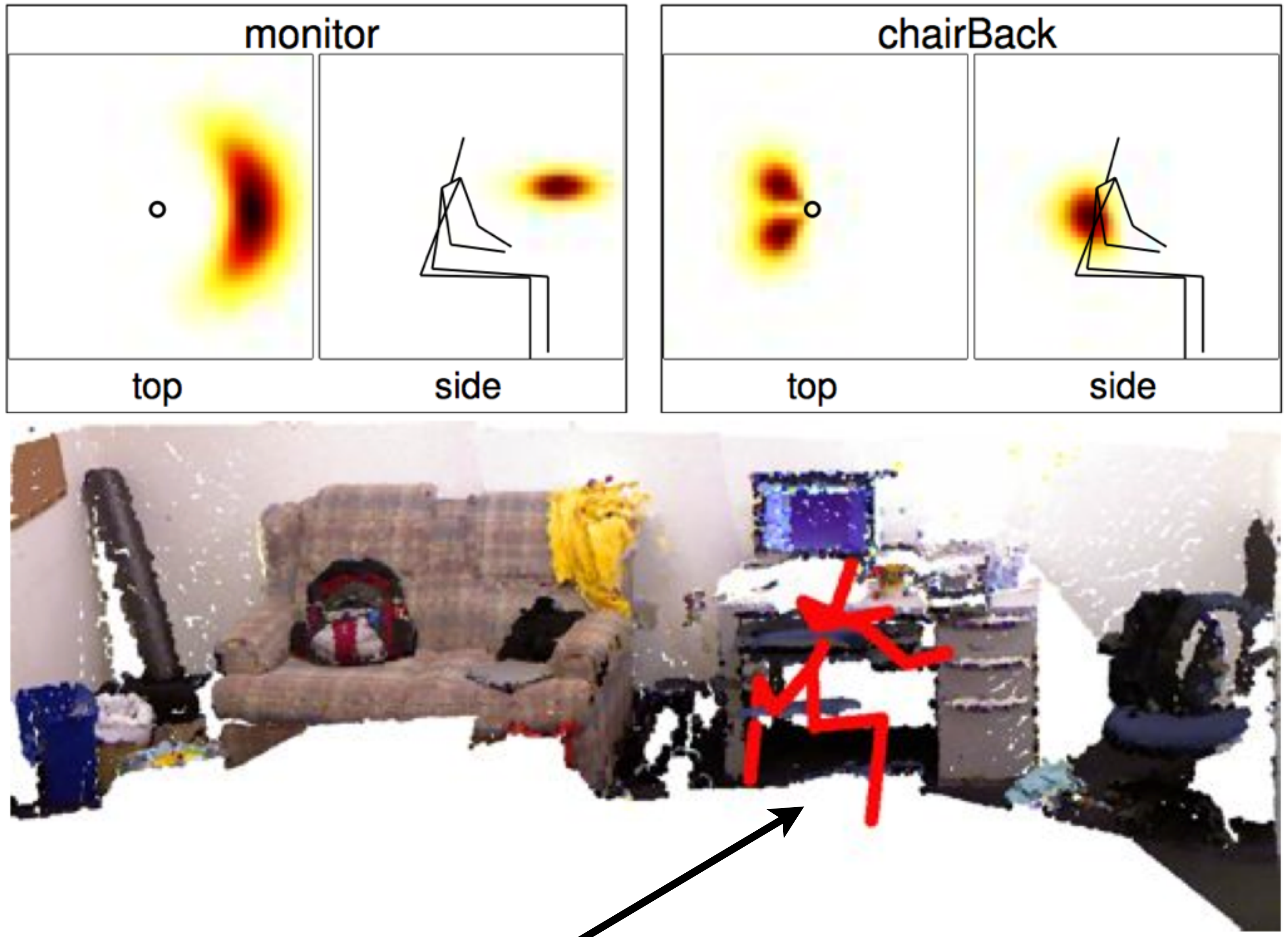
Affordance is

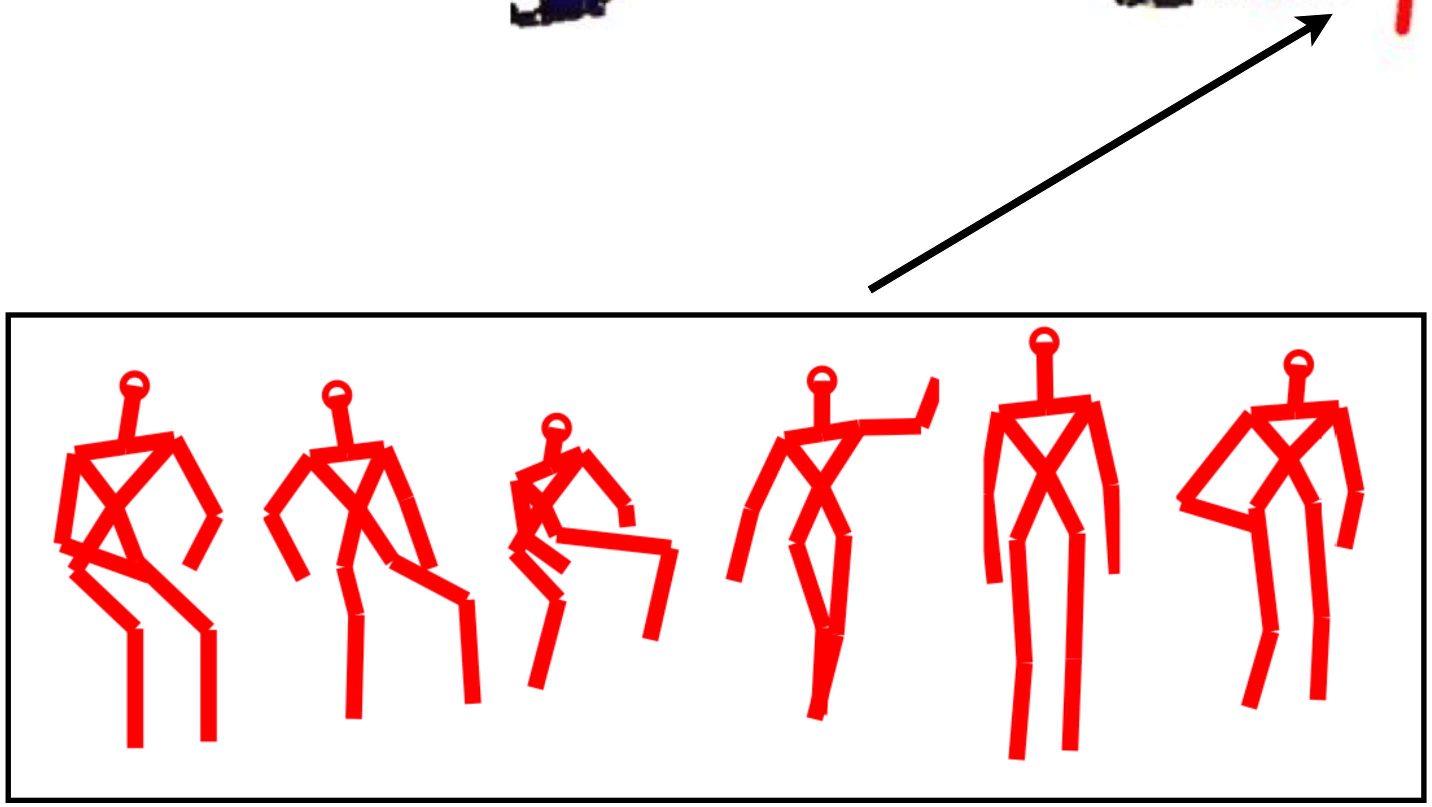
Other Potential Applications:

- **Populating Virtual Environments**
- **Interaction-aware Design**
- **Functional Understanding**



Previous Work: Hallucinating People

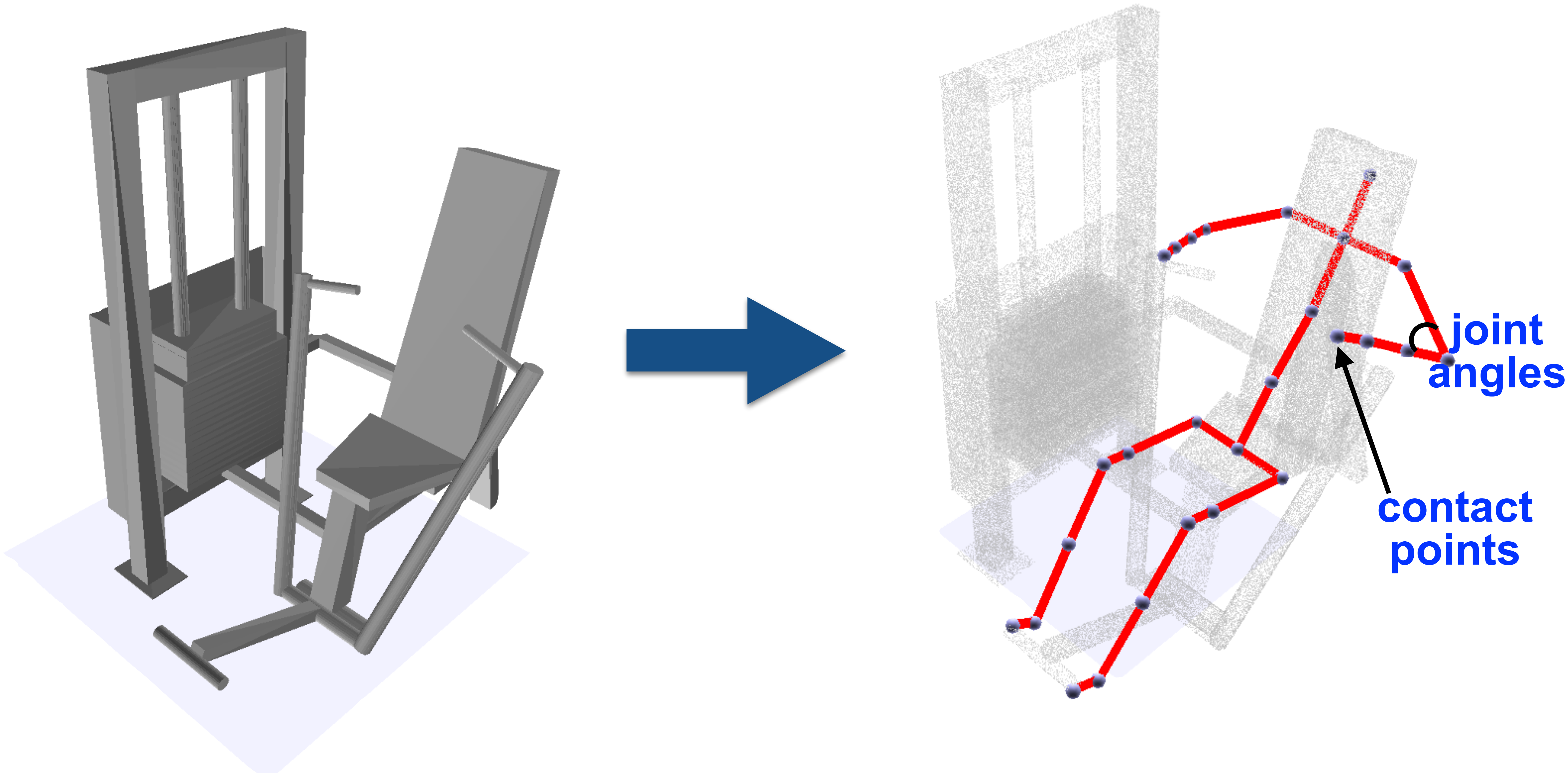




Jiang'13
Gupta et al.'11
Grabner et al.'11

Goal

Predict an arbitrary pose



Overview

Introduction

Learning Affordance Model

Pose Prediction

Results & Applications

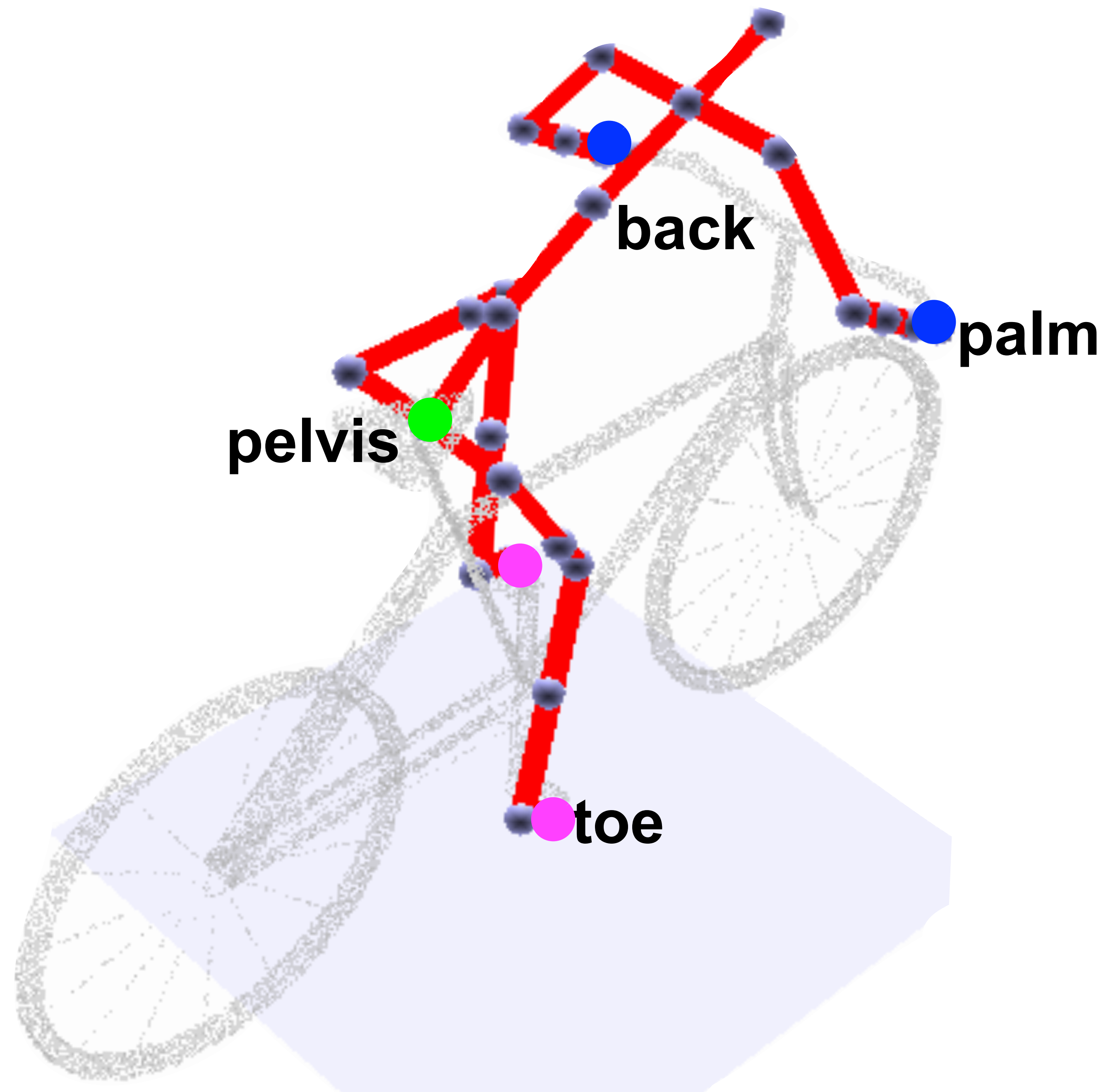
Affordance Model

Pose Parameters

- ➡ Contact points m
- ○ Joint Angles θ, T

$$m : P \rightarrow S$$

Body parts Shape



Affordance Model

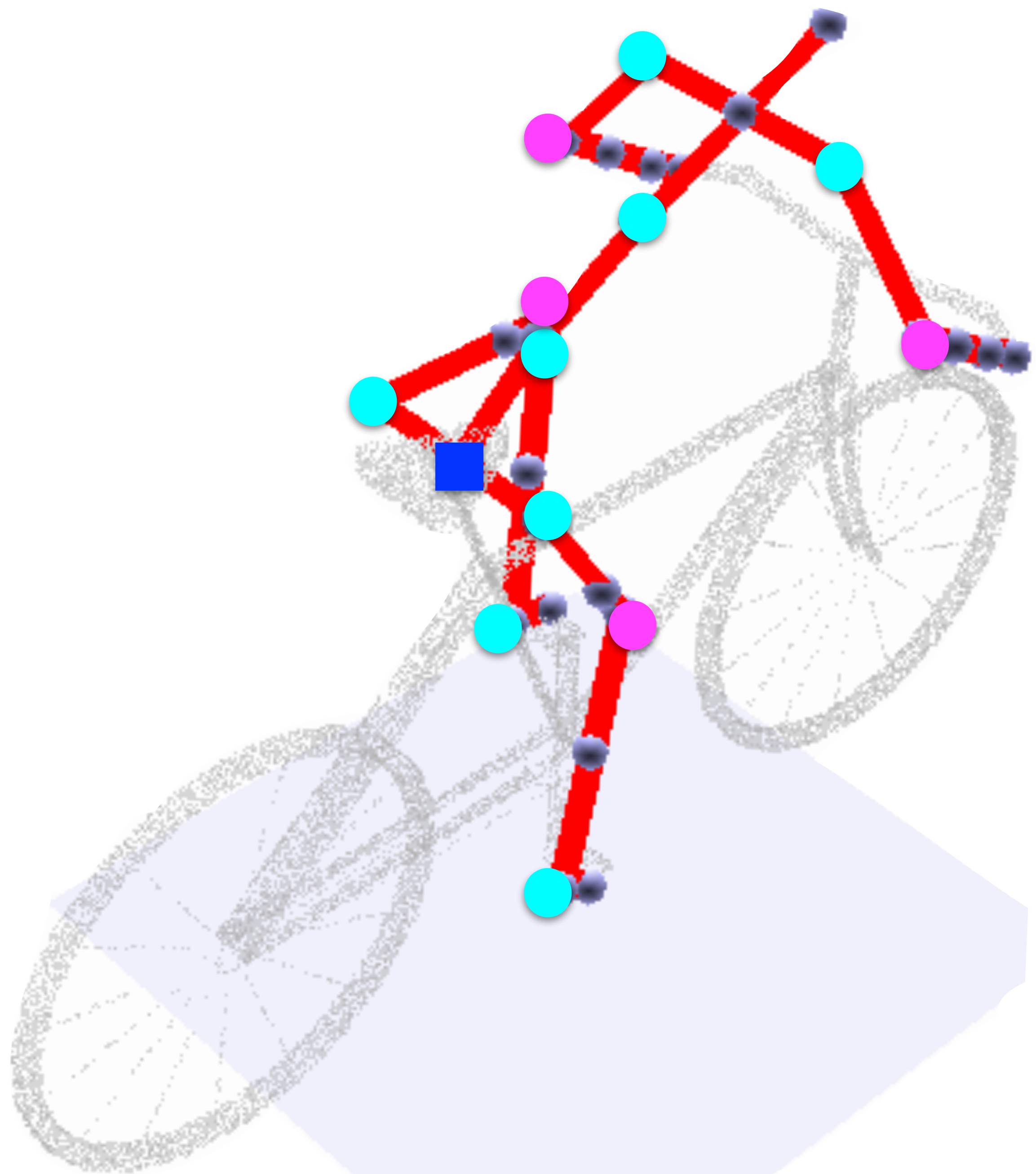
Pose Parameters

- Contact points m
- ➡ Joint Angles θ, T

3DOF + T (Ball & socket reference joint)

3DOF (Ball & socket joint)

1DOF (Hinge joint)



Affordance Model

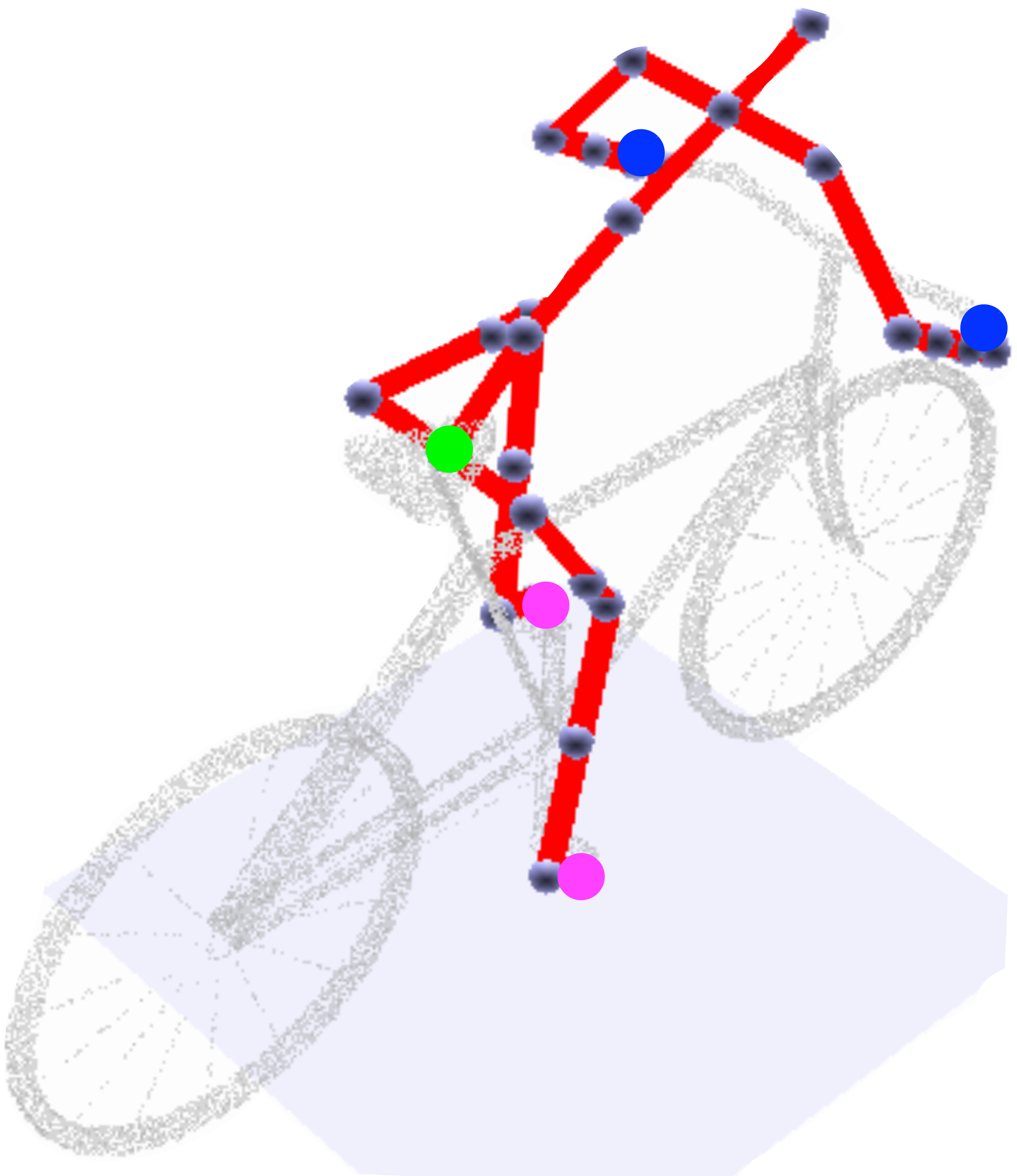
Pose Parameters

- Contact points m
- Joint Angles θ, T

$$E_{\text{dist}} = \sum_{p \in P} ||T\mathbf{p}_\theta - m(p)||^2$$

Energy

- ➡ Contact Distance
- Feature Compatibility
- Pose Prior
- Symmetry
- Surface Intersections



Affordance Model

Pose Parameters

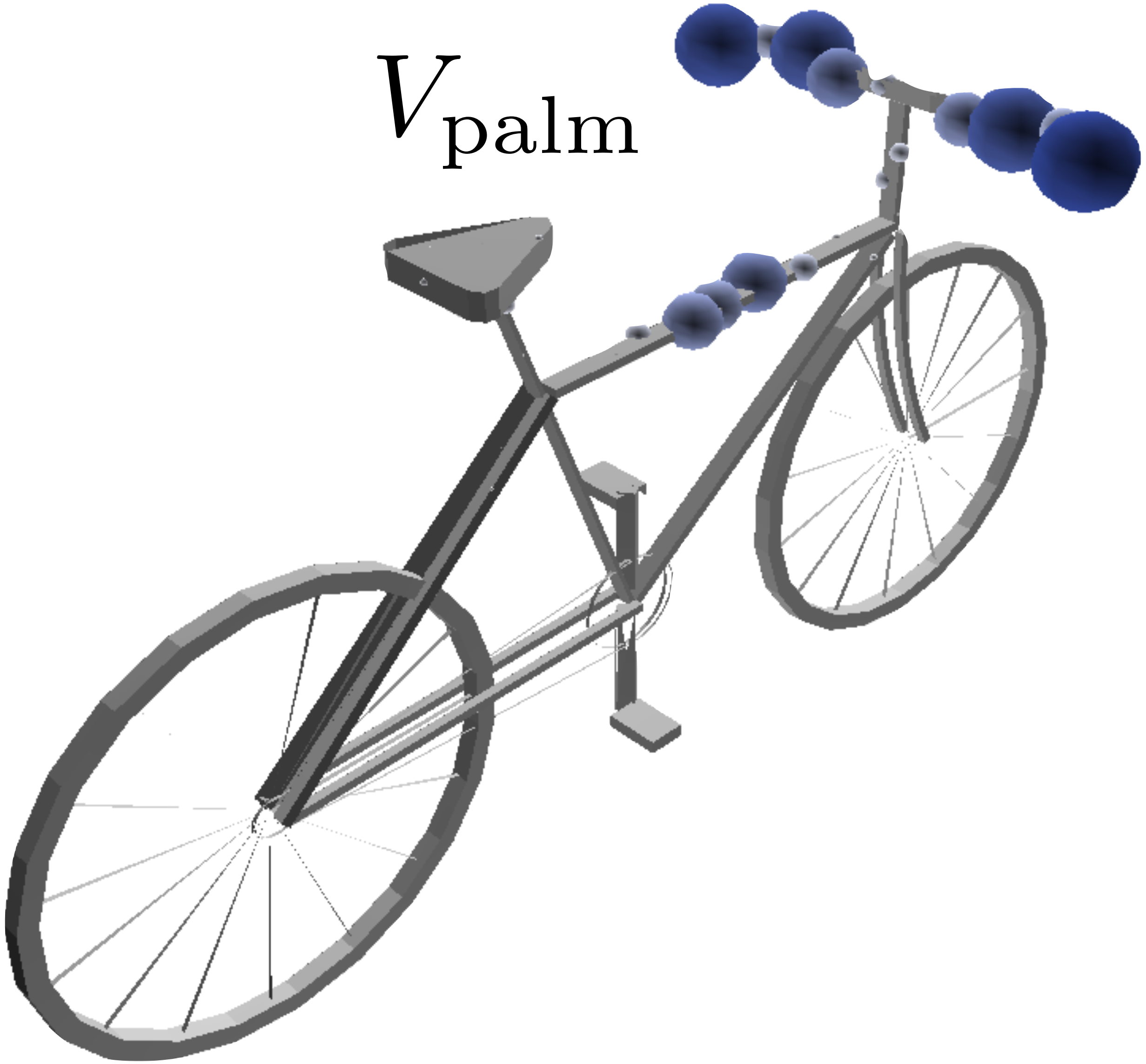
- Contact points m
- Joint Angles θ, T

$$E_{\text{feat}} = \sum_{p \in P} -\log V_p(m(p))$$

Trained Classifier

Energy

- Contact Distance
- ➡ Feature Compatibility
- Pose Prior
- Symmetry
- Surface Intersections



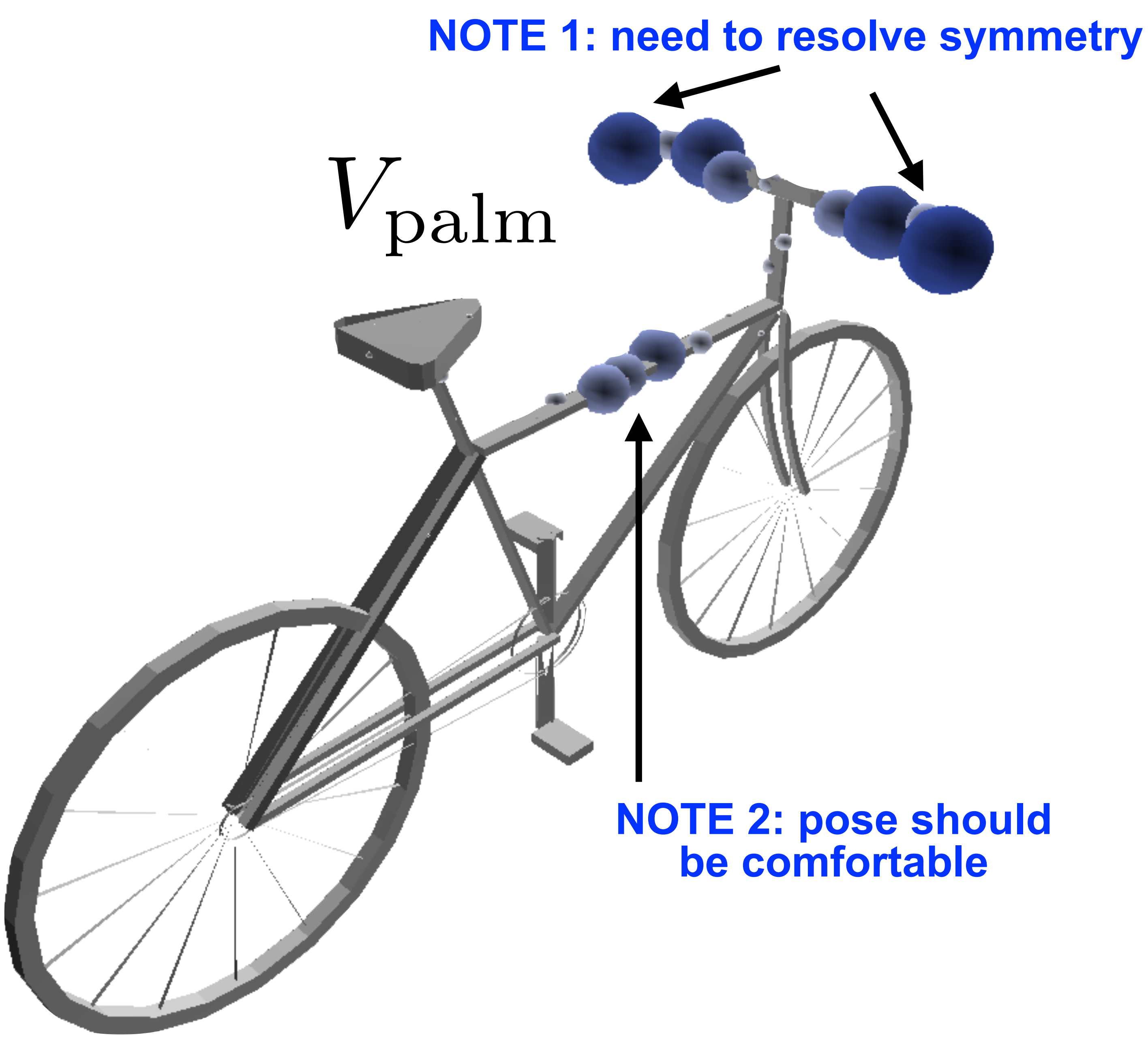
Affordance Model

Pose Parameters

- Contact points m
- Joint Angles θ, T

$$E_{\text{feat}} = \sum_{p \in P} -\log V_p(m(p))$$

Trained Classifier



Energy

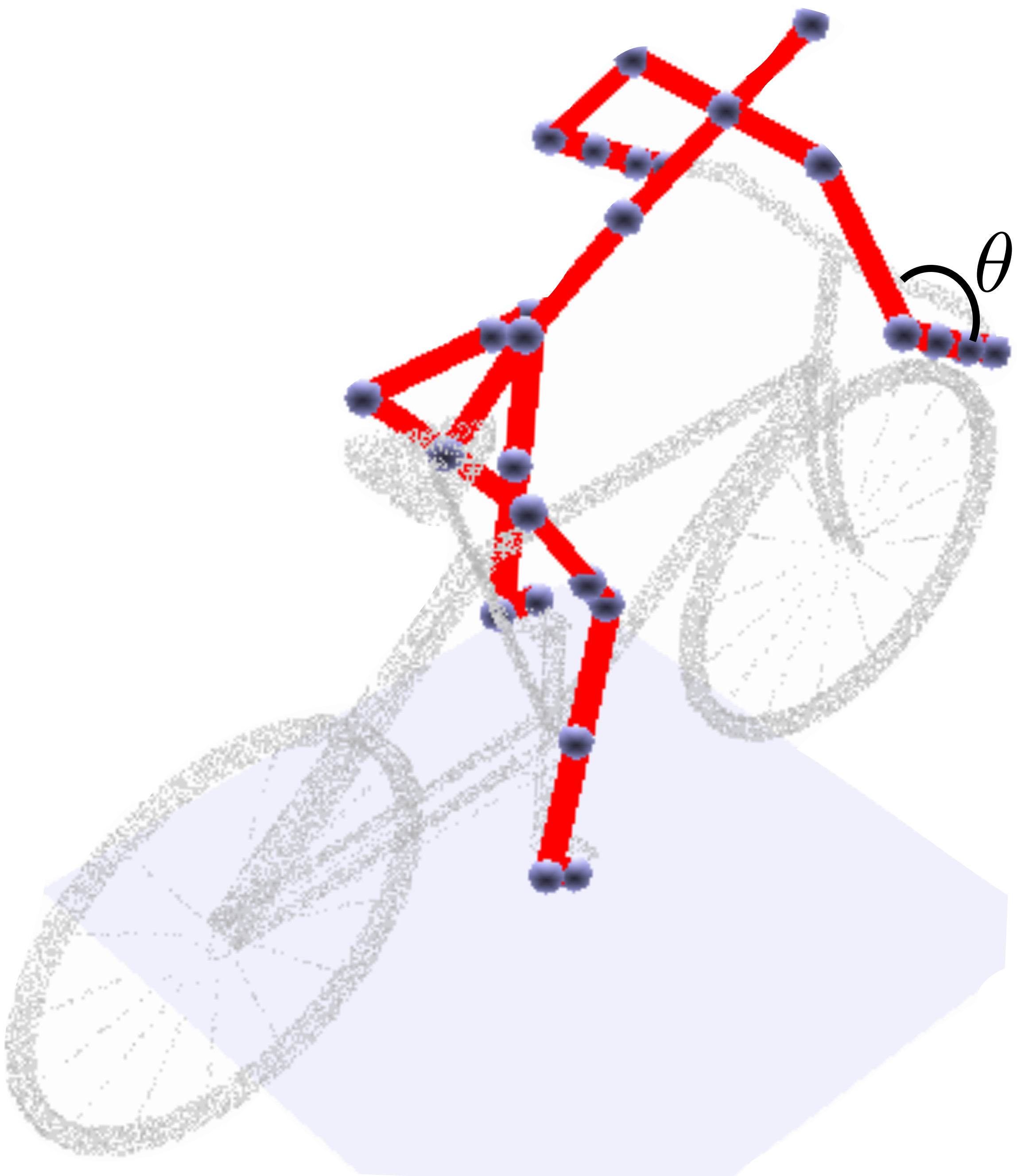
- Contact Distance
- ➡ Feature Compatibility
- Pose Prior
- Symmetry
- Surface Intersections

Affordance Model

$$E_{\text{pose}} = \min_{l \in L} \sum_{i}^{40} \frac{|\theta_i - \mu_l^l|^2}{(\sigma_i^l)^2}$$

Pose Parameters

- Contact points m
- Joint Angles θ, T



Energy

- Contact Distance
- Feature Compatibility
- ➡ Pose Prior
- Symmetry
- Surface Intersections

Affordance Model

Pose Parameters

- Contact points m
- Joint Angles θ, T

Energy

- ➡ Contact Distance
- ➡ Feature Compatibility
- ➡ Pose Prior
- Symmetry
- Surface Intersections

Hard Constraint

Key Optimization Terms

Additional Terms

Overview

Introduction

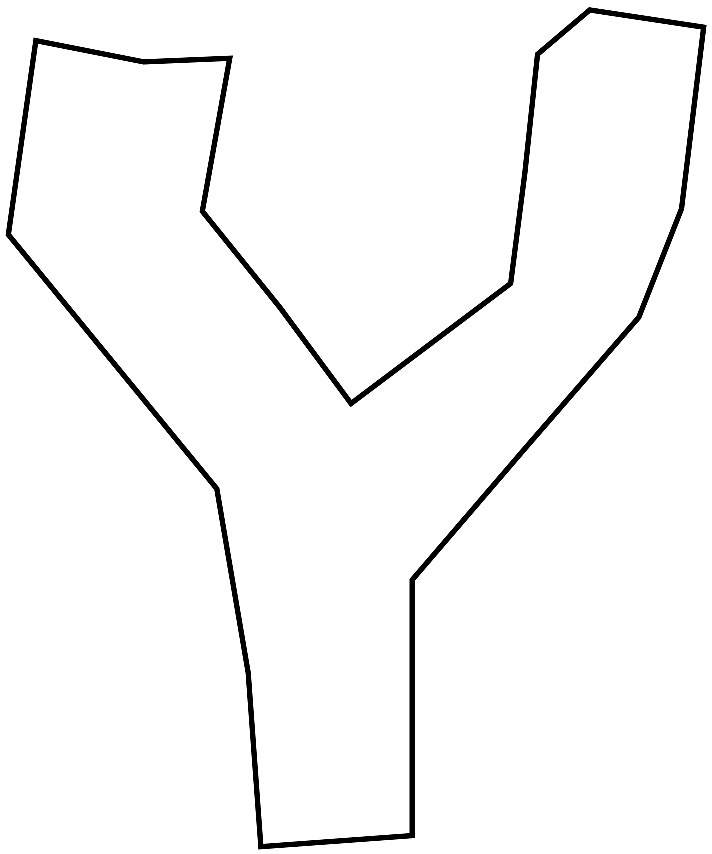
Learning Affordance Model

Pose Prediction

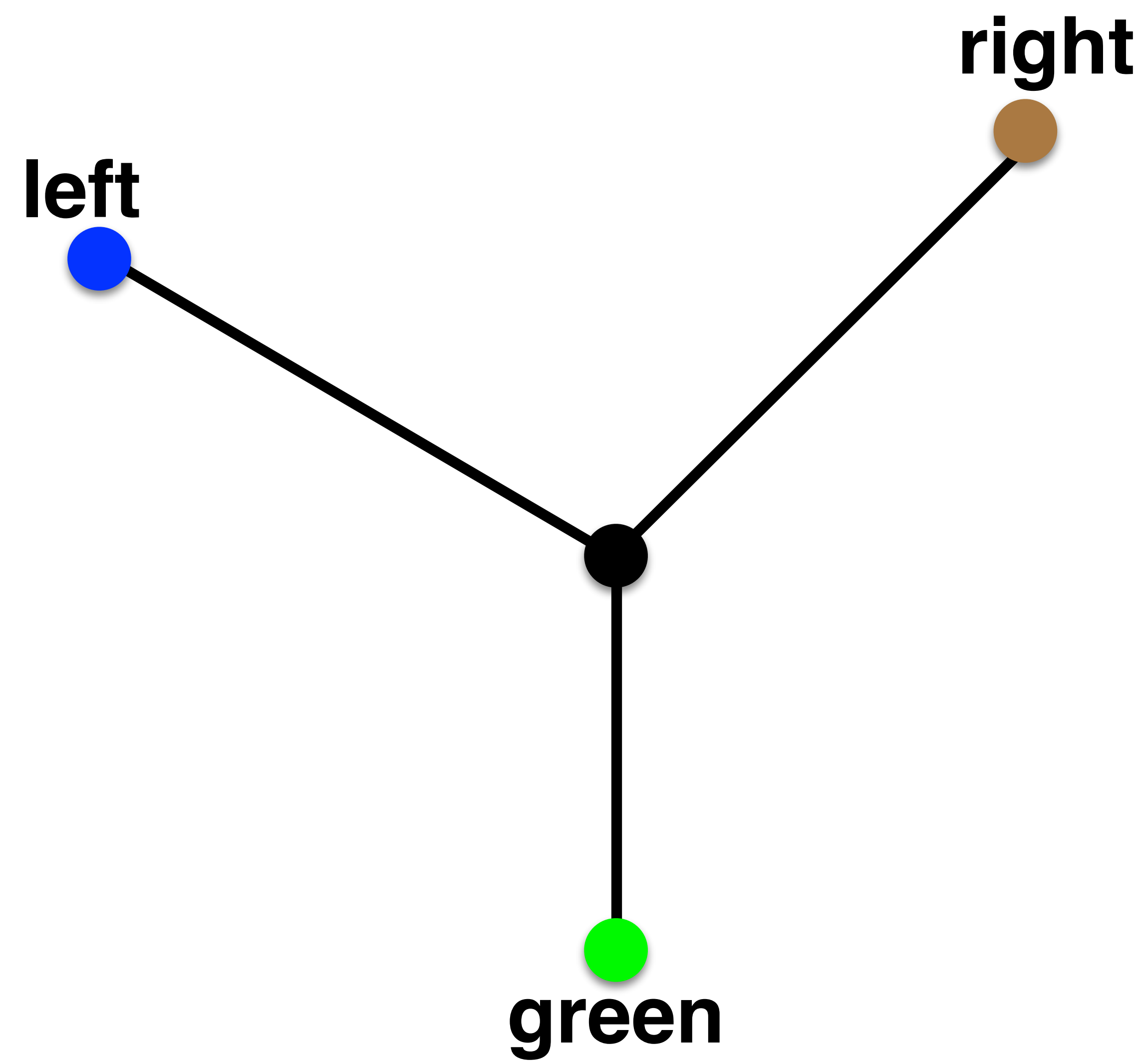
Results & Applications

Pose Prediction

Input



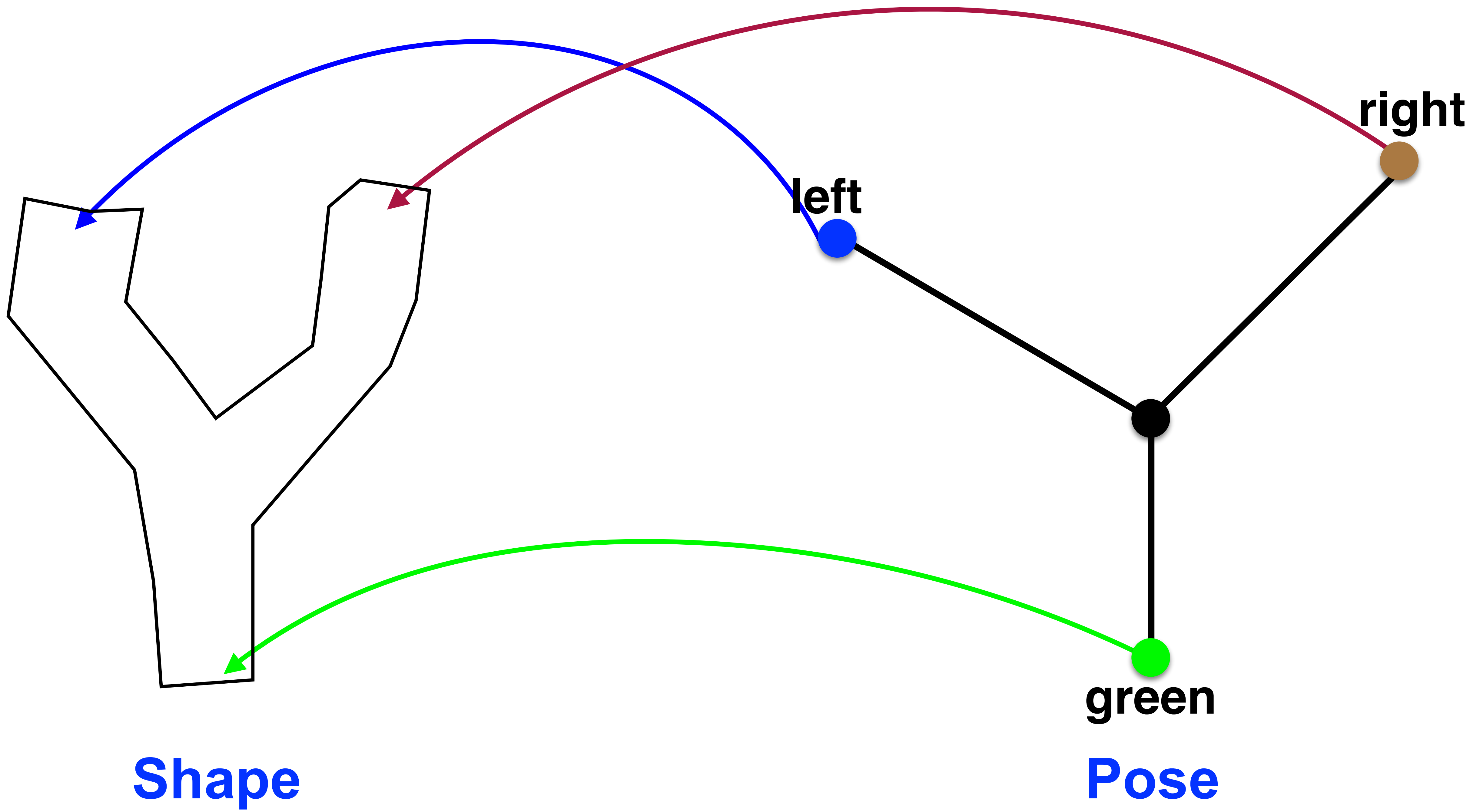
Shape



Pose

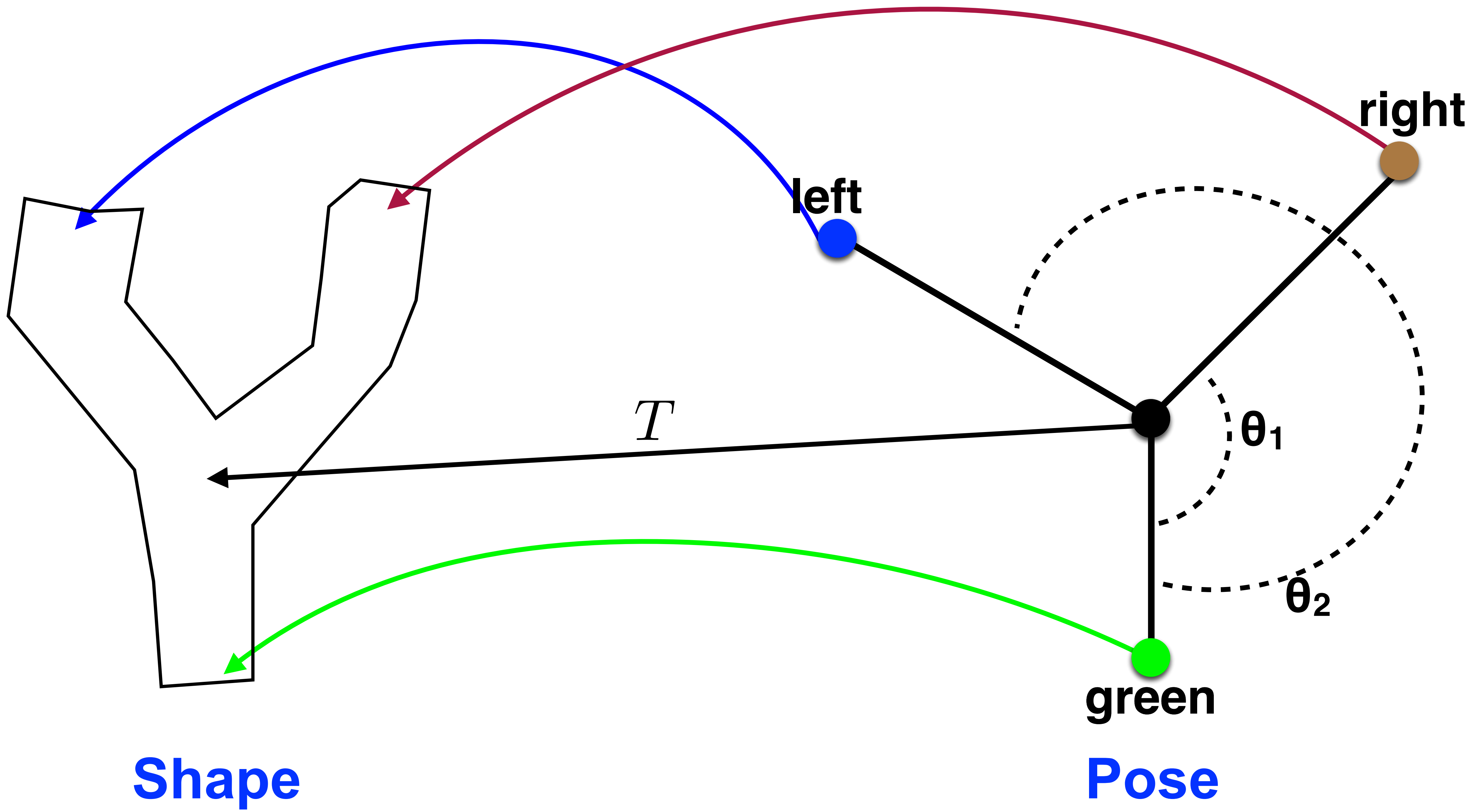
Pose Prediction

Output: m



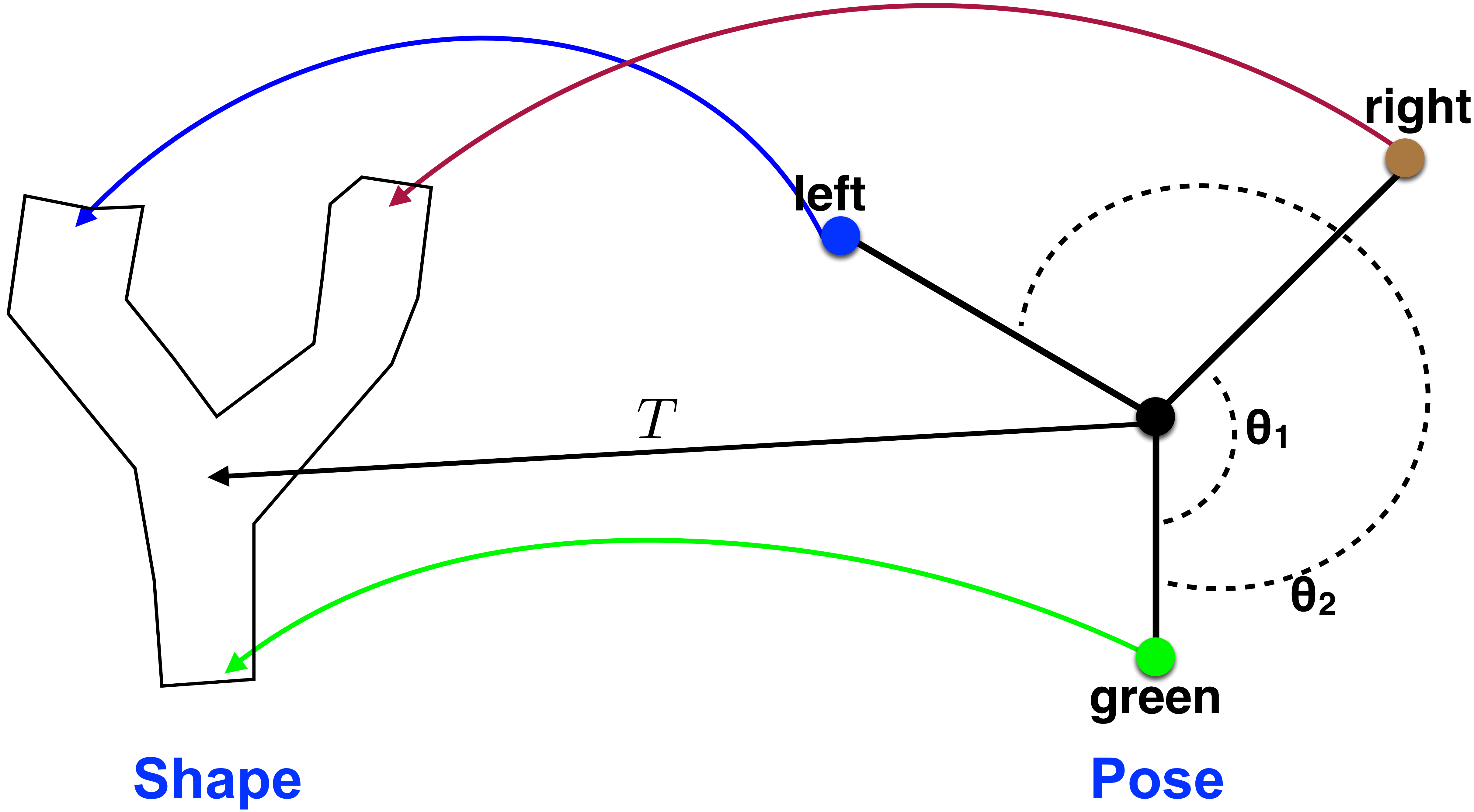
Pose Prediction

Output: m θ, T



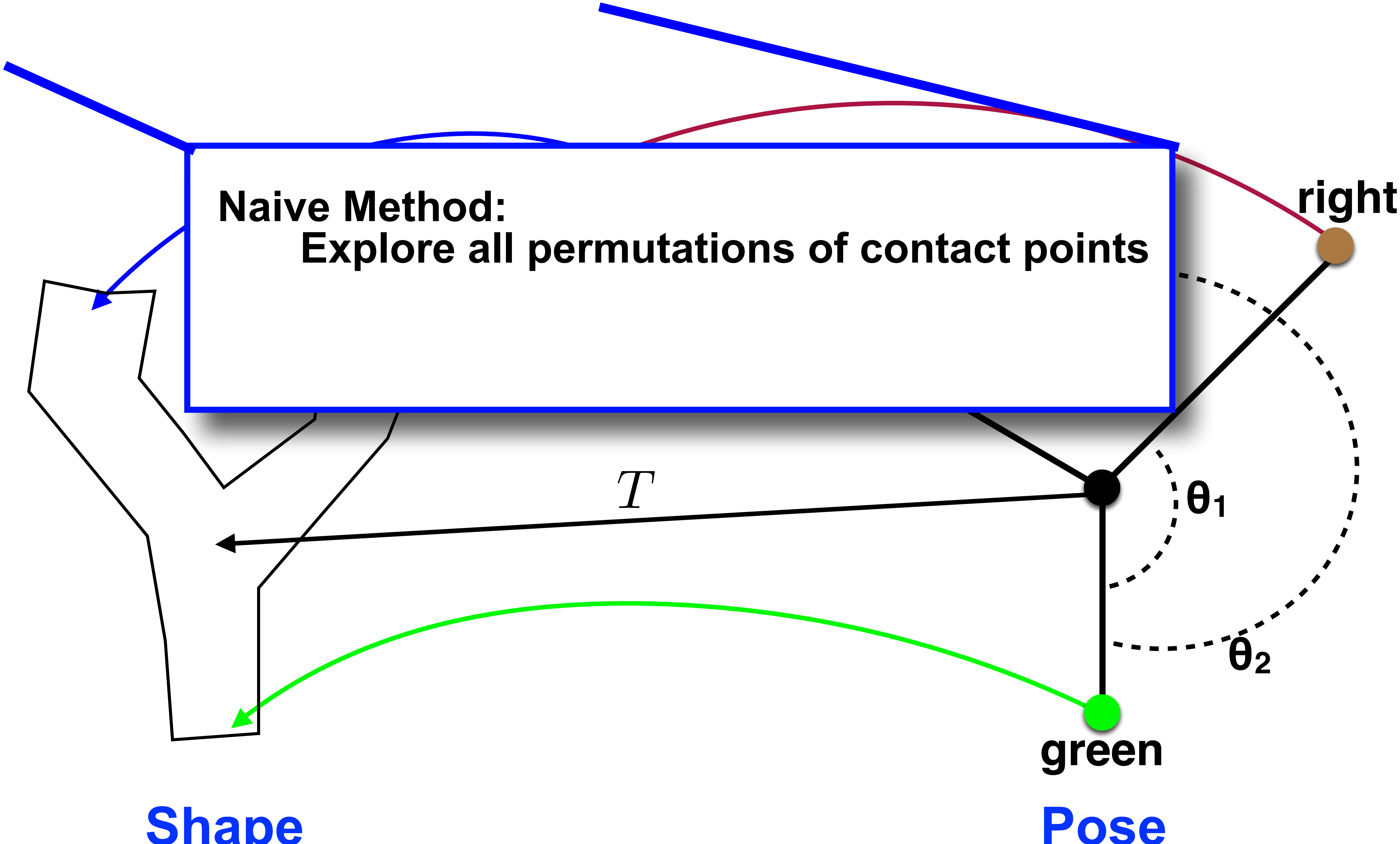
Pose Prediction

Output: $m \xrightarrow[\text{IK}]{} \theta, T$



Pose Prediction

Output: $m \xrightarrow[\text{IK}]{} \theta, T$

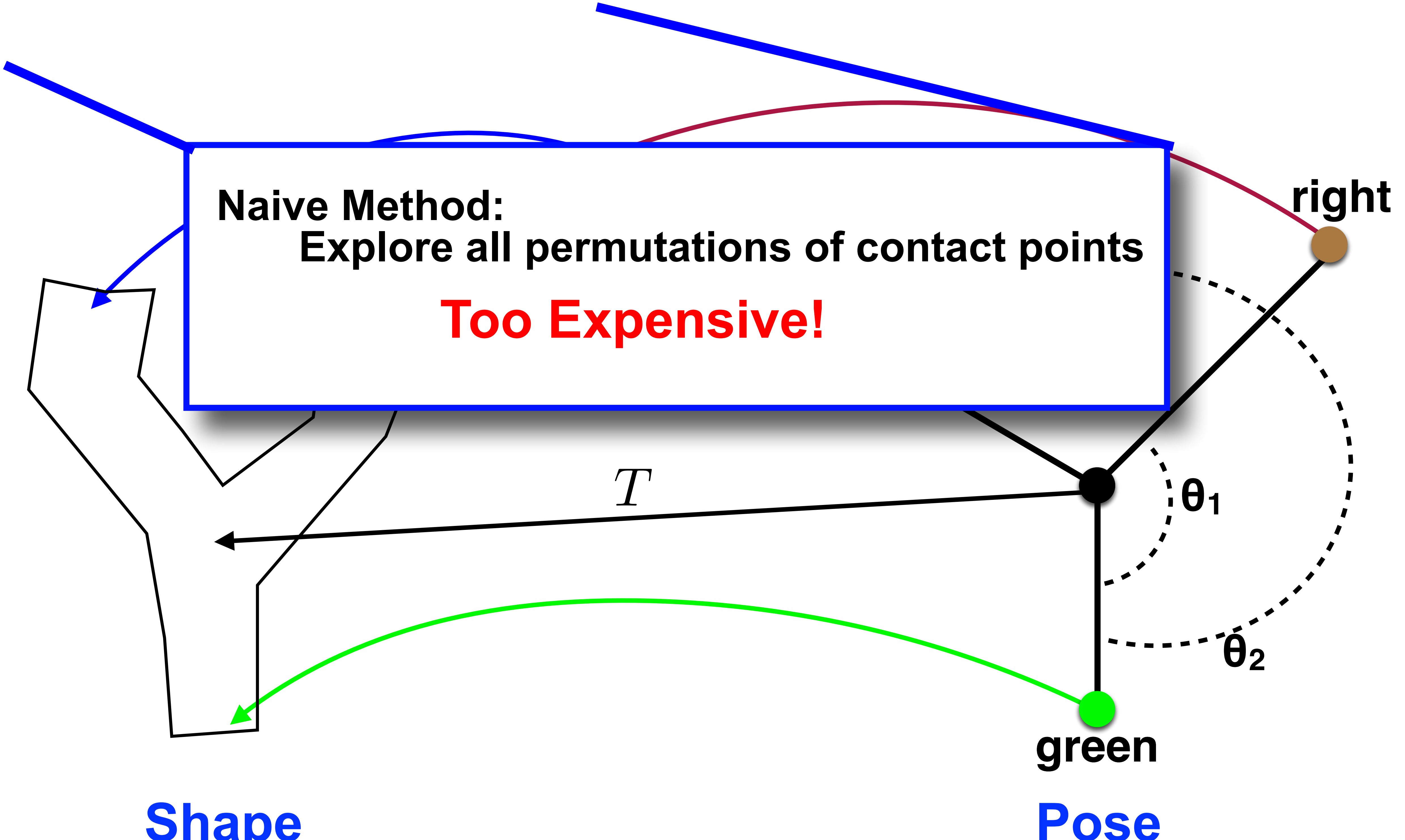


Pose Prediction

Output: $m \xrightarrow[\text{IK}]{} \theta, T$

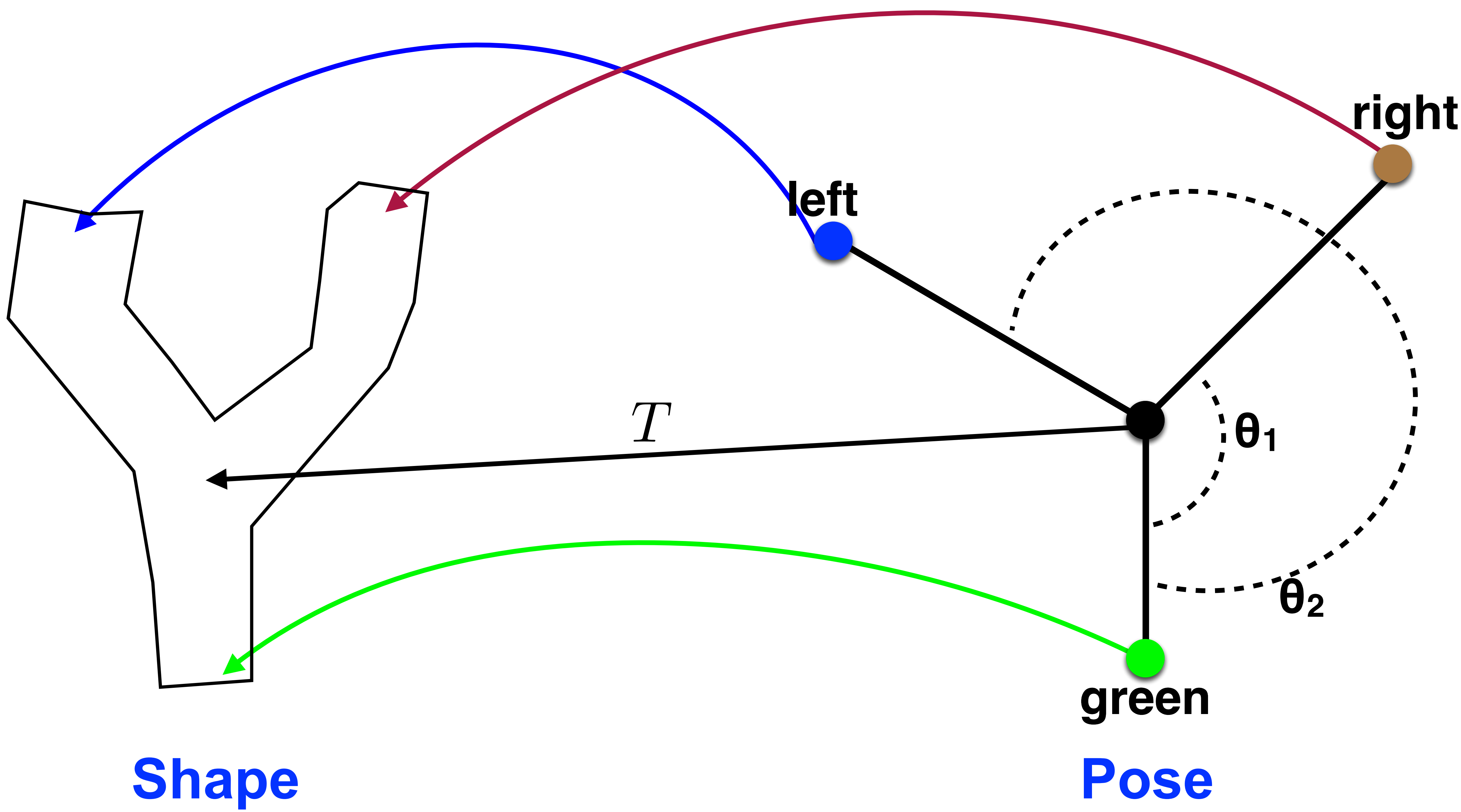
Naive Method:
Explore all permutations of contact points

Too Expensive!



Pose Prediction

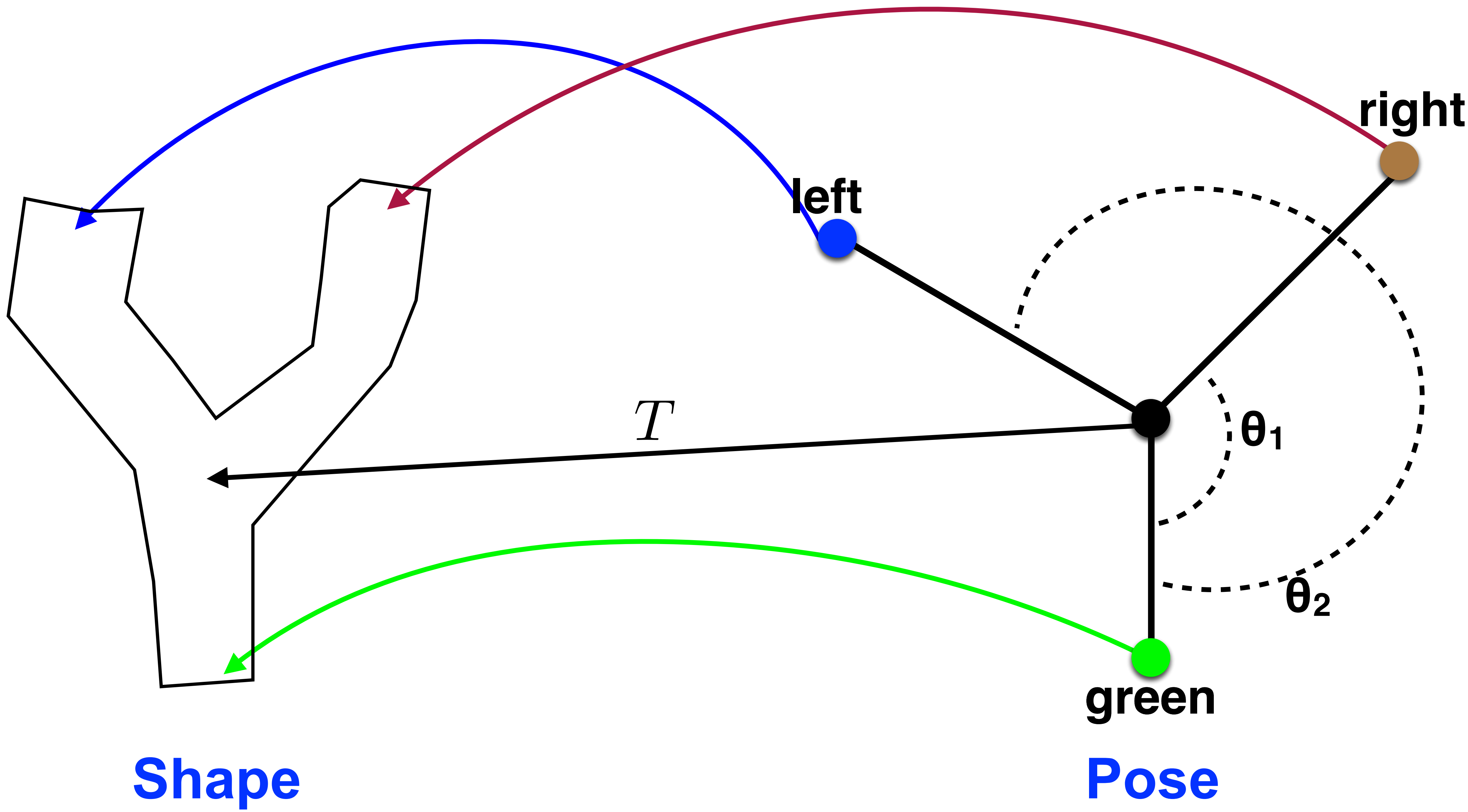
Output: $m \xleftarrow{\text{FK}} \theta, T$



Pose Prediction

Output: $\underline{m} \longleftrightarrow \underline{\theta, T}$

Contact Distribution — **Pose Prior**



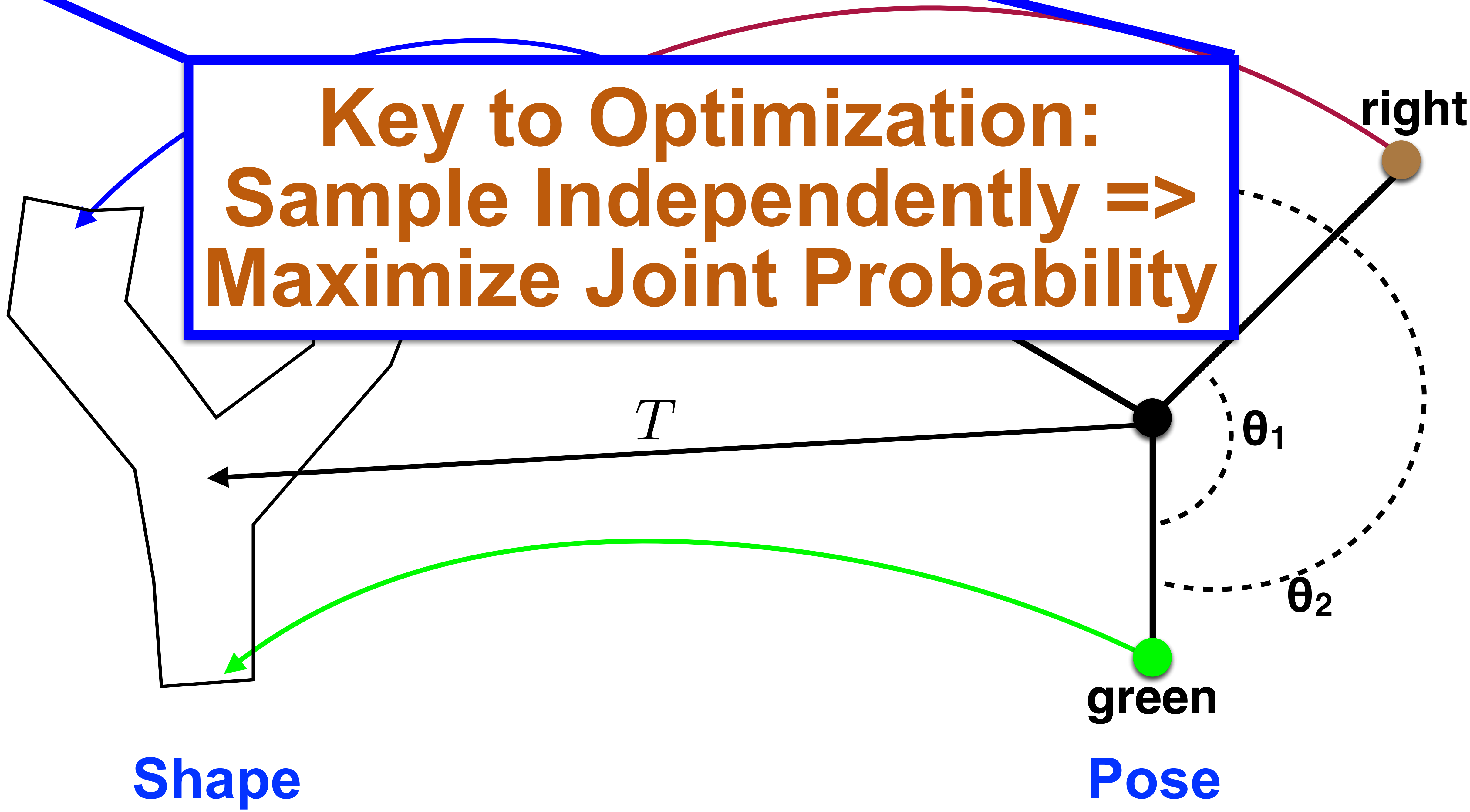
Pose Prediction

Output: $m \longleftrightarrow \theta, T$

Contact Distribution

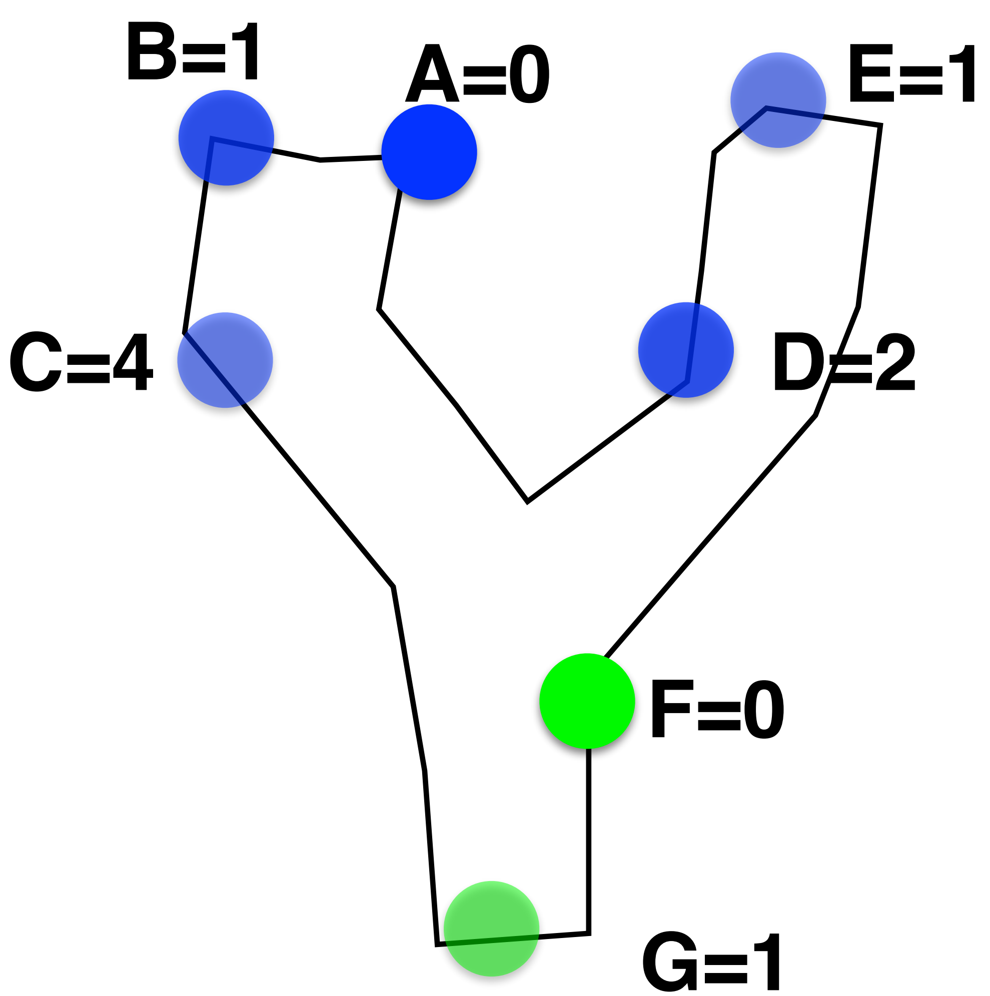
Pose Prior

Key to Optimization: Sample Independently => Maximize Joint Probability

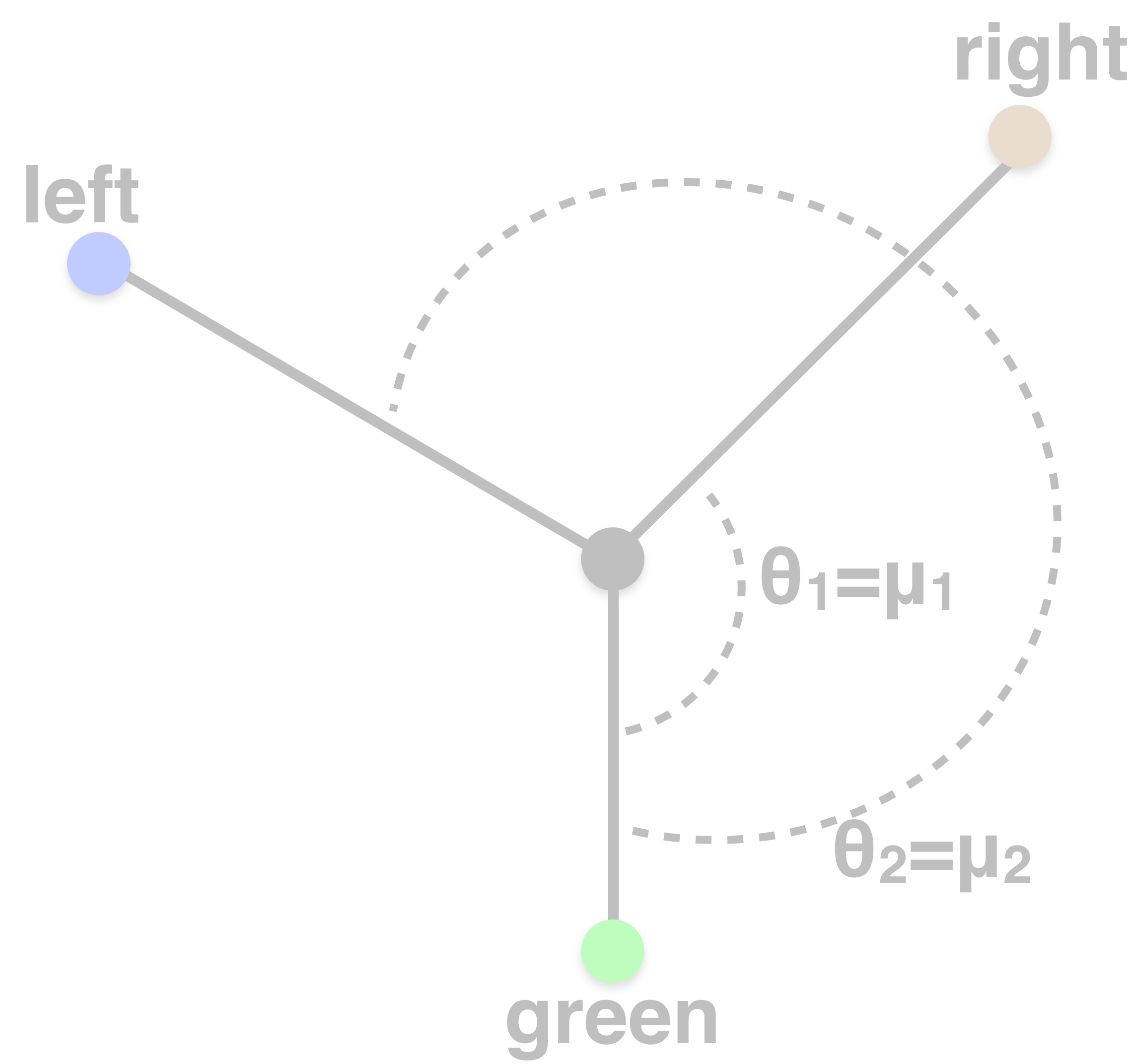


Pose Prediction

Sample m: classify surface based on local features



Contact Distribution

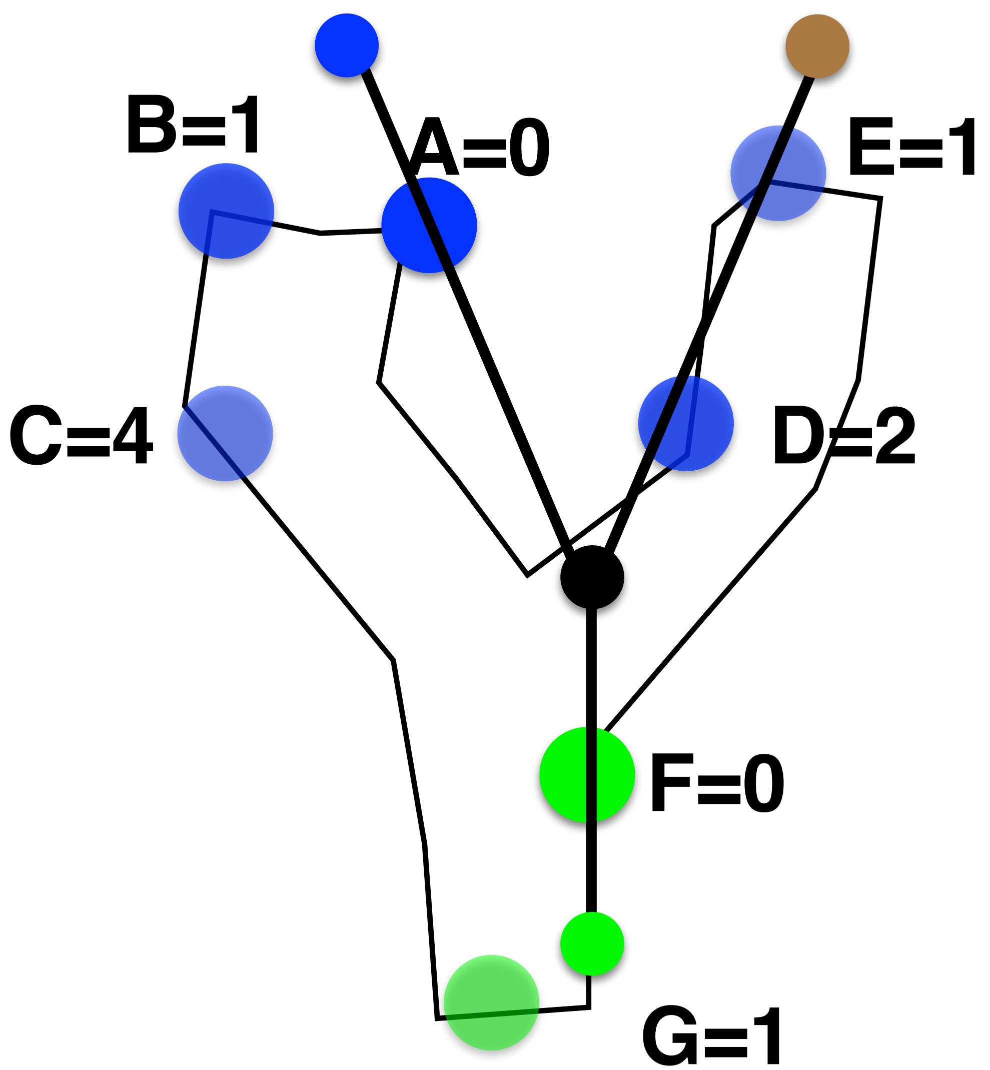


Pose

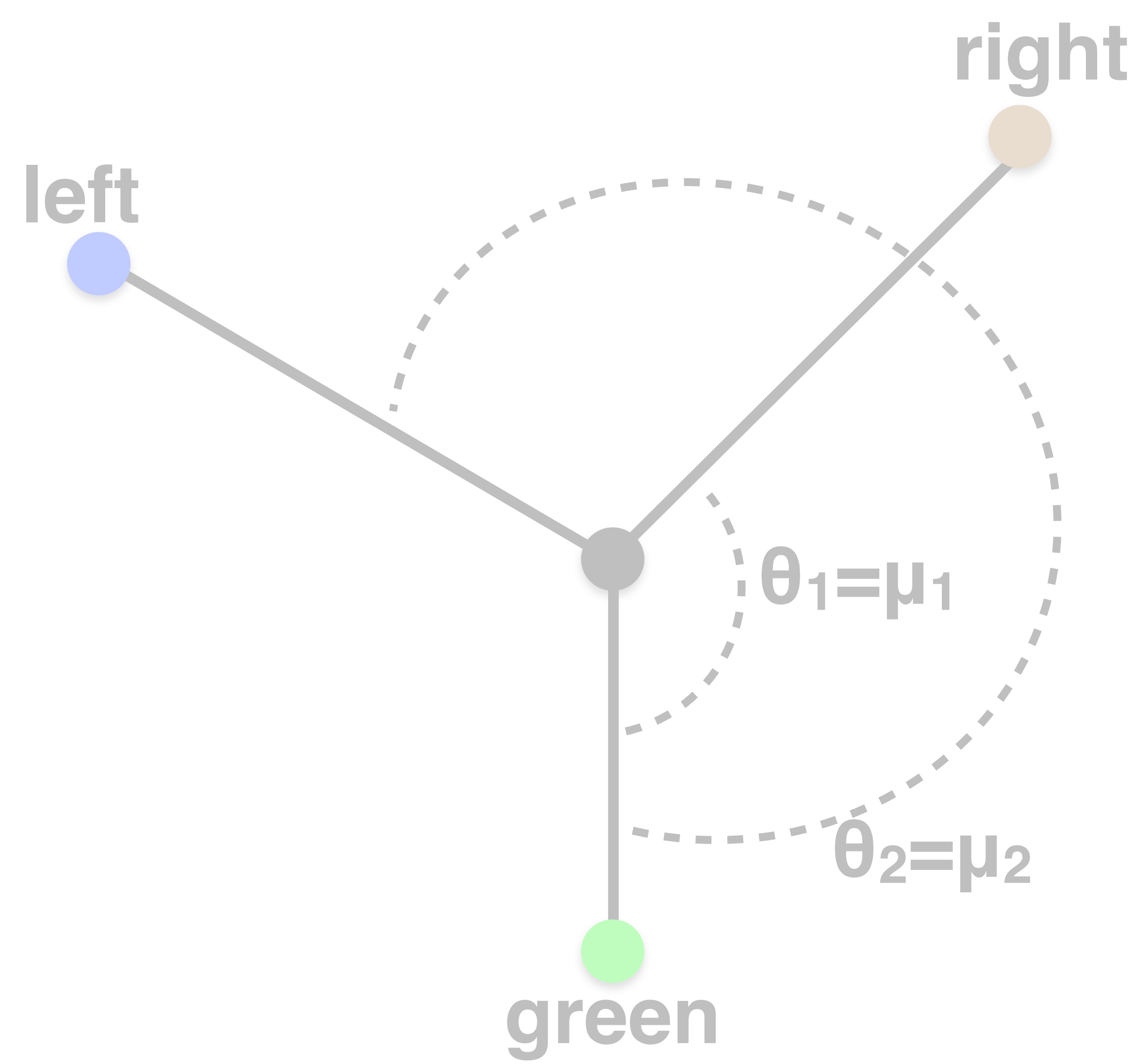
Pose Prediction

Sample m: classify surface based on local features

Need to include the pose prior in optimization!!!



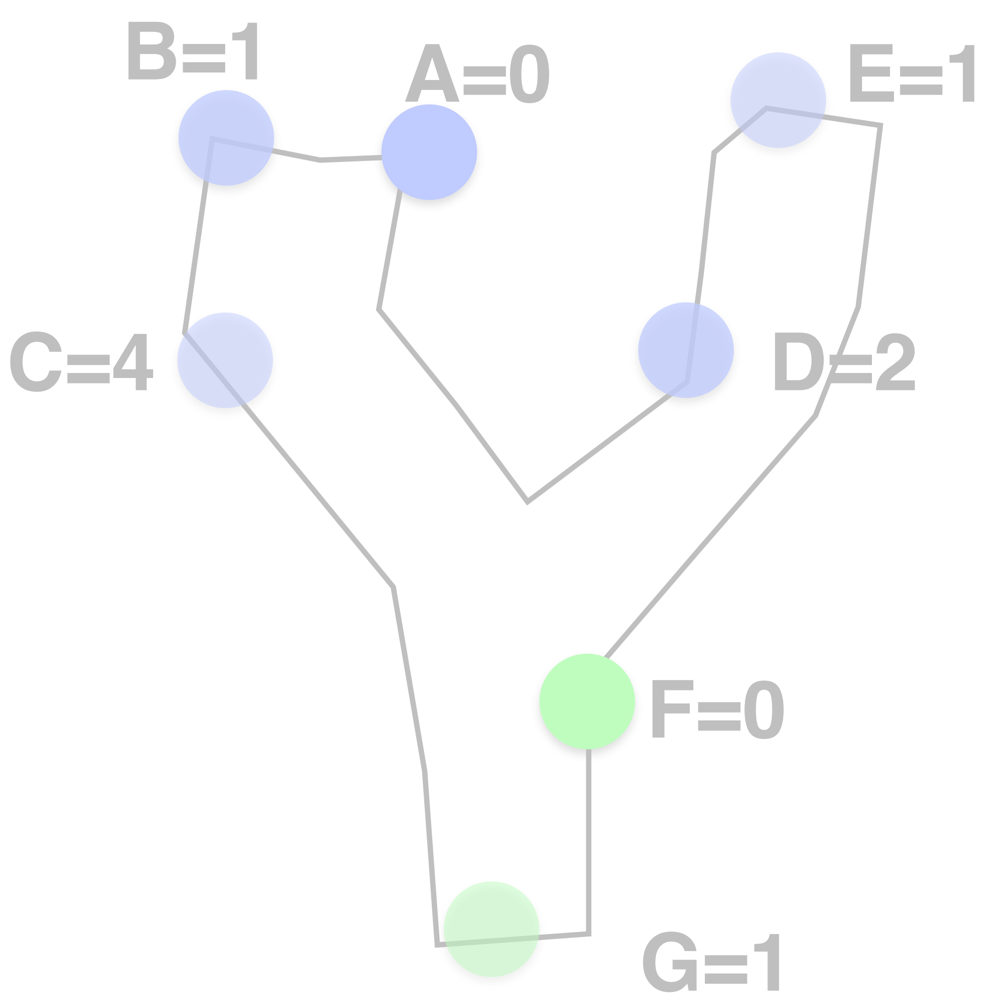
Contact Distribution



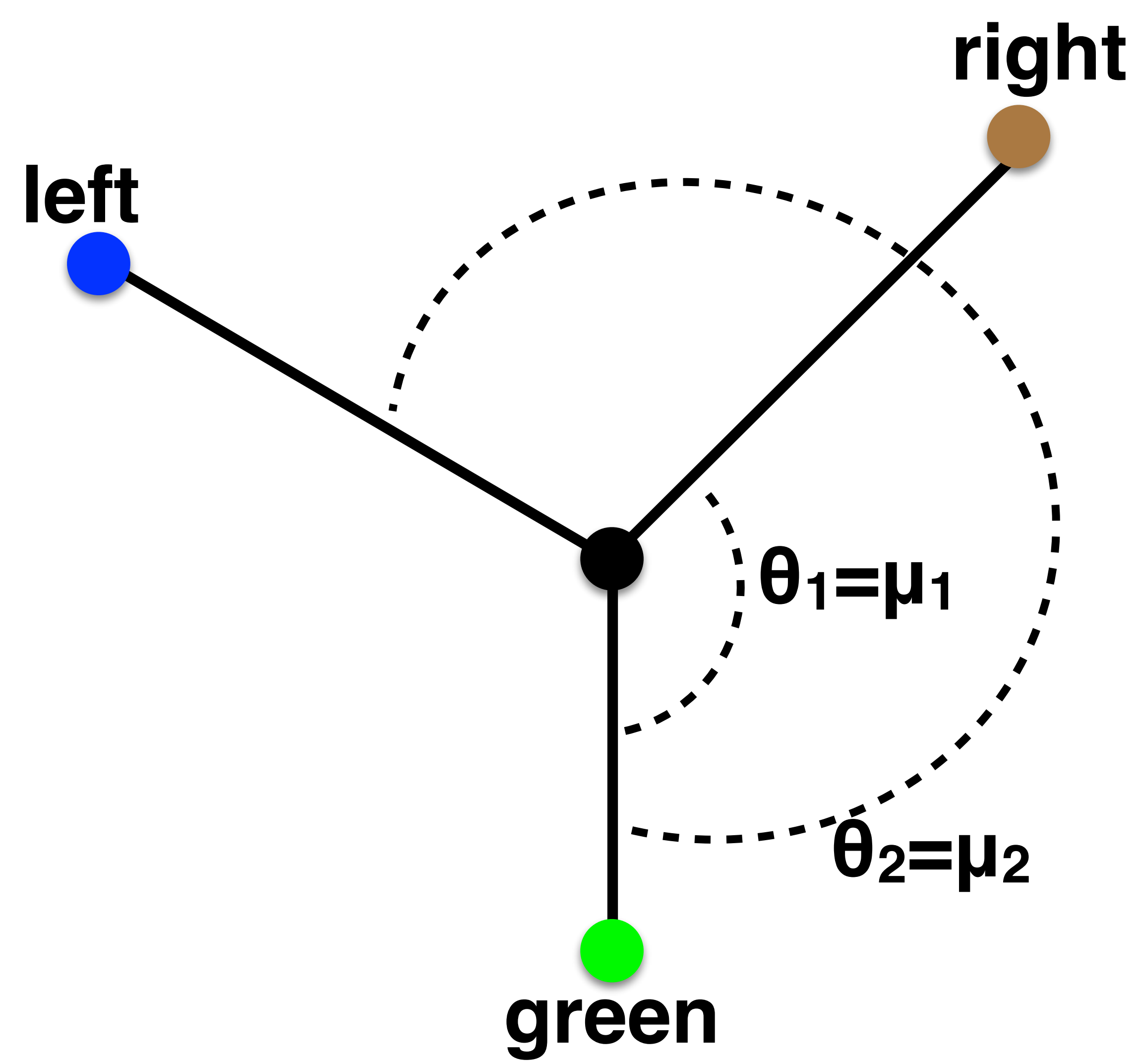
Pose

Pose Prediction

Sample θ, T: pose is represented by two Gaussians



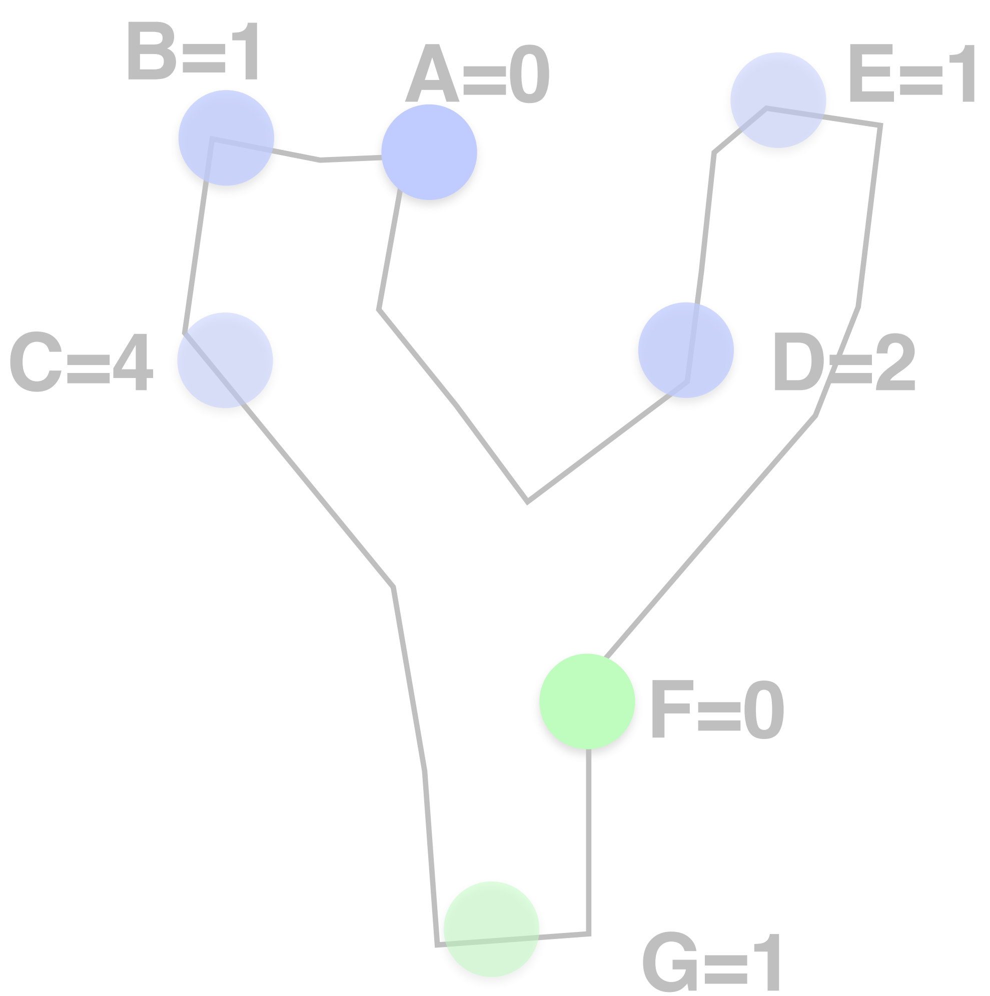
Contact Distribution



Pose

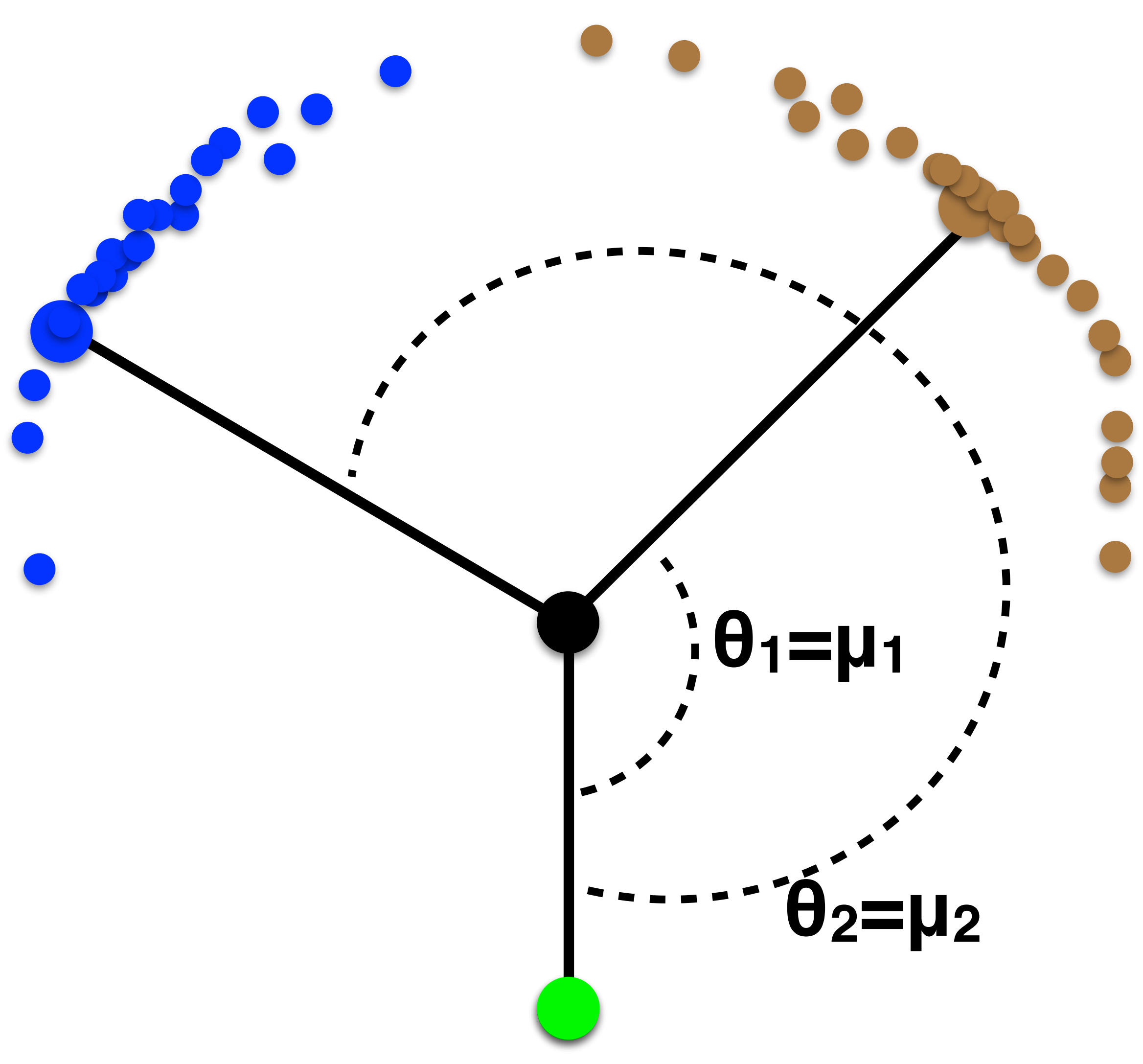
Pose Prediction

Sample θ, T: pose is represented by two Gaussians



Contact Distribution

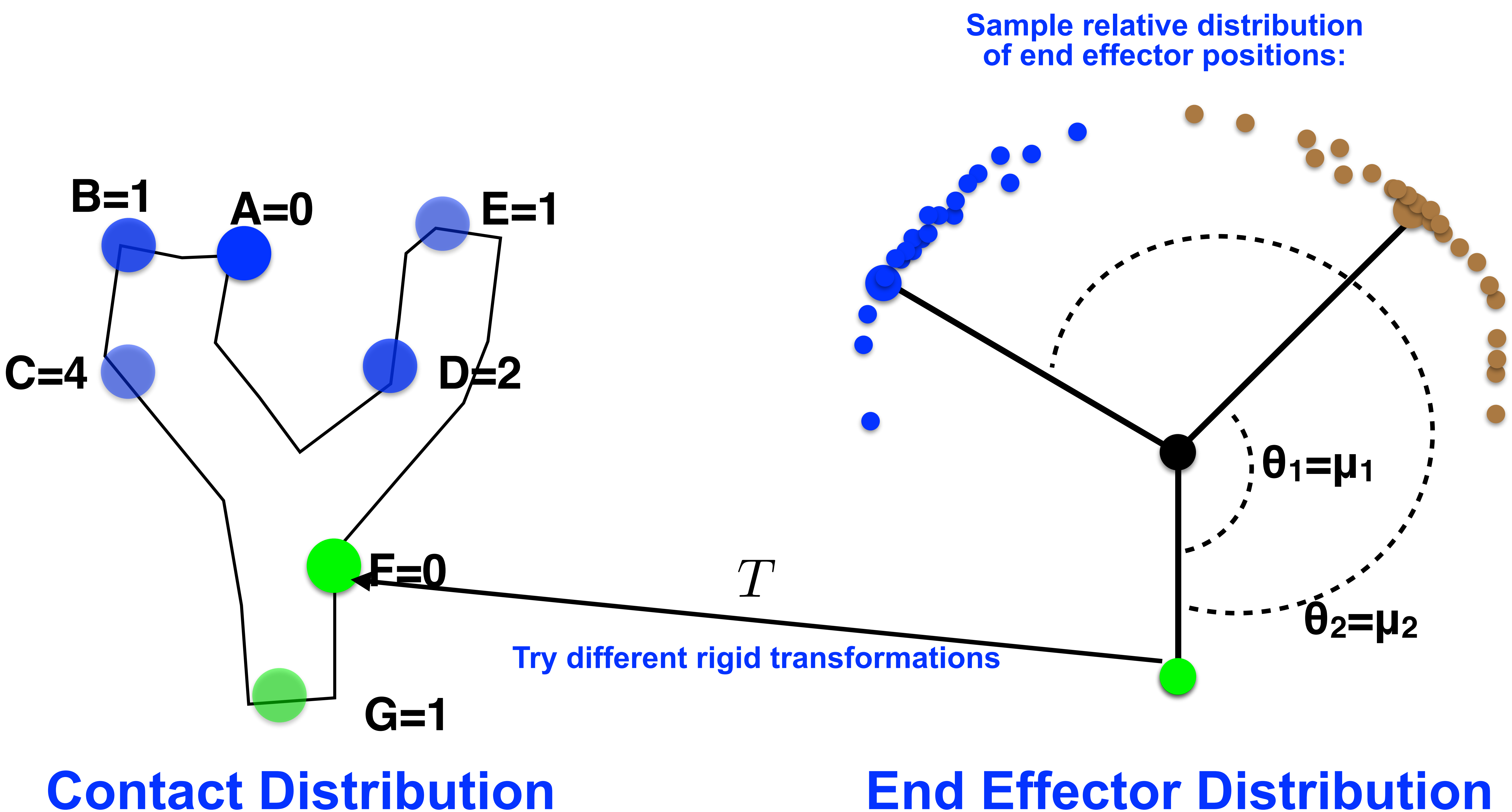
Sample relative distribution of end effector positions:



End Effector Distribution

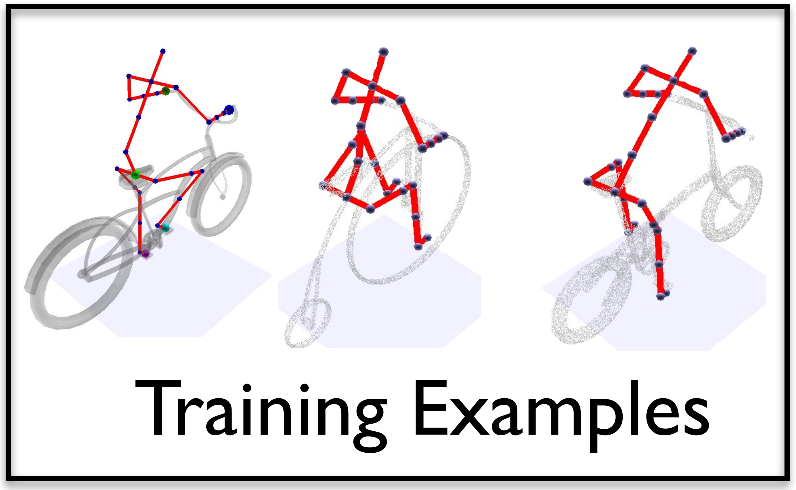
Pose Prediction

Sample θ, T: pose is represented by two Gaussians



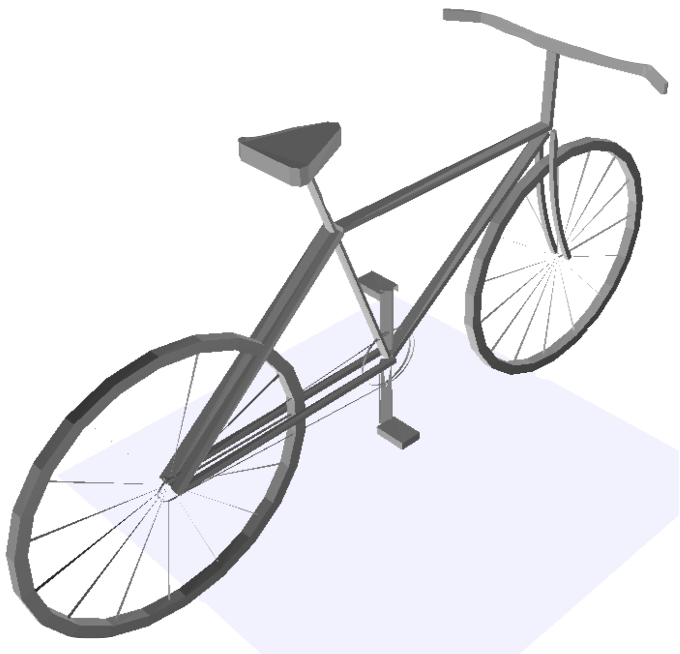
Pose Prediction

Our pipeline

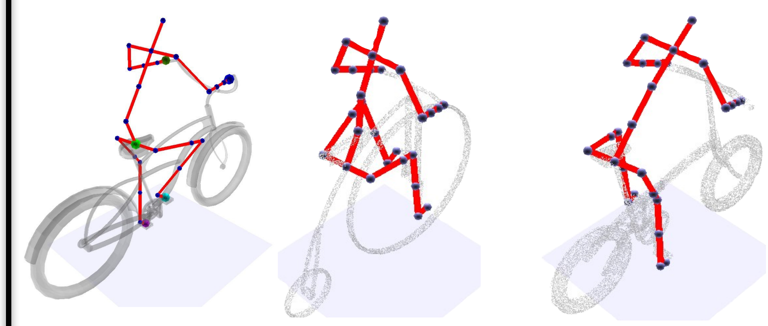


Pose Prediction

Our pipeline



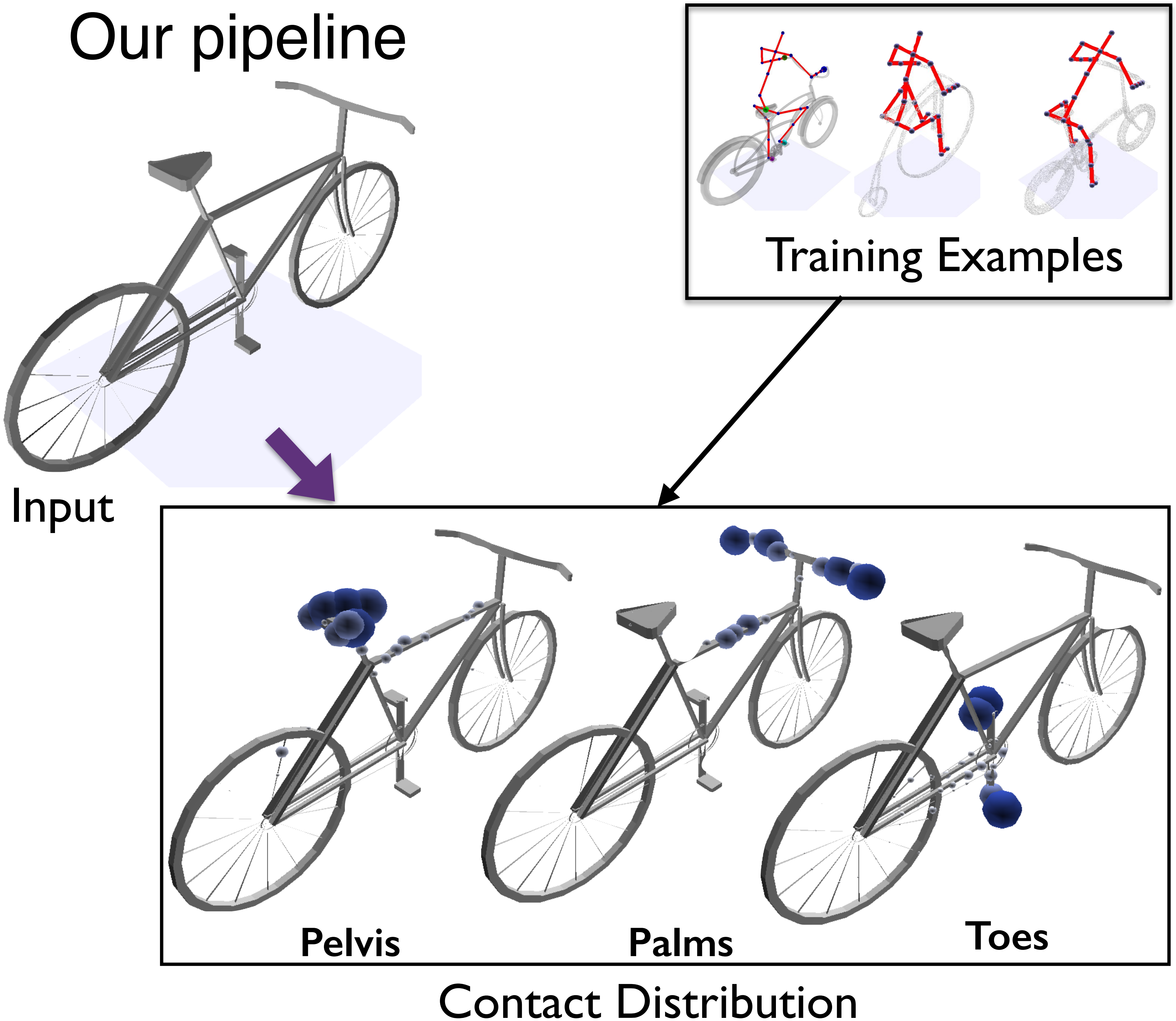
Input



Training Examples

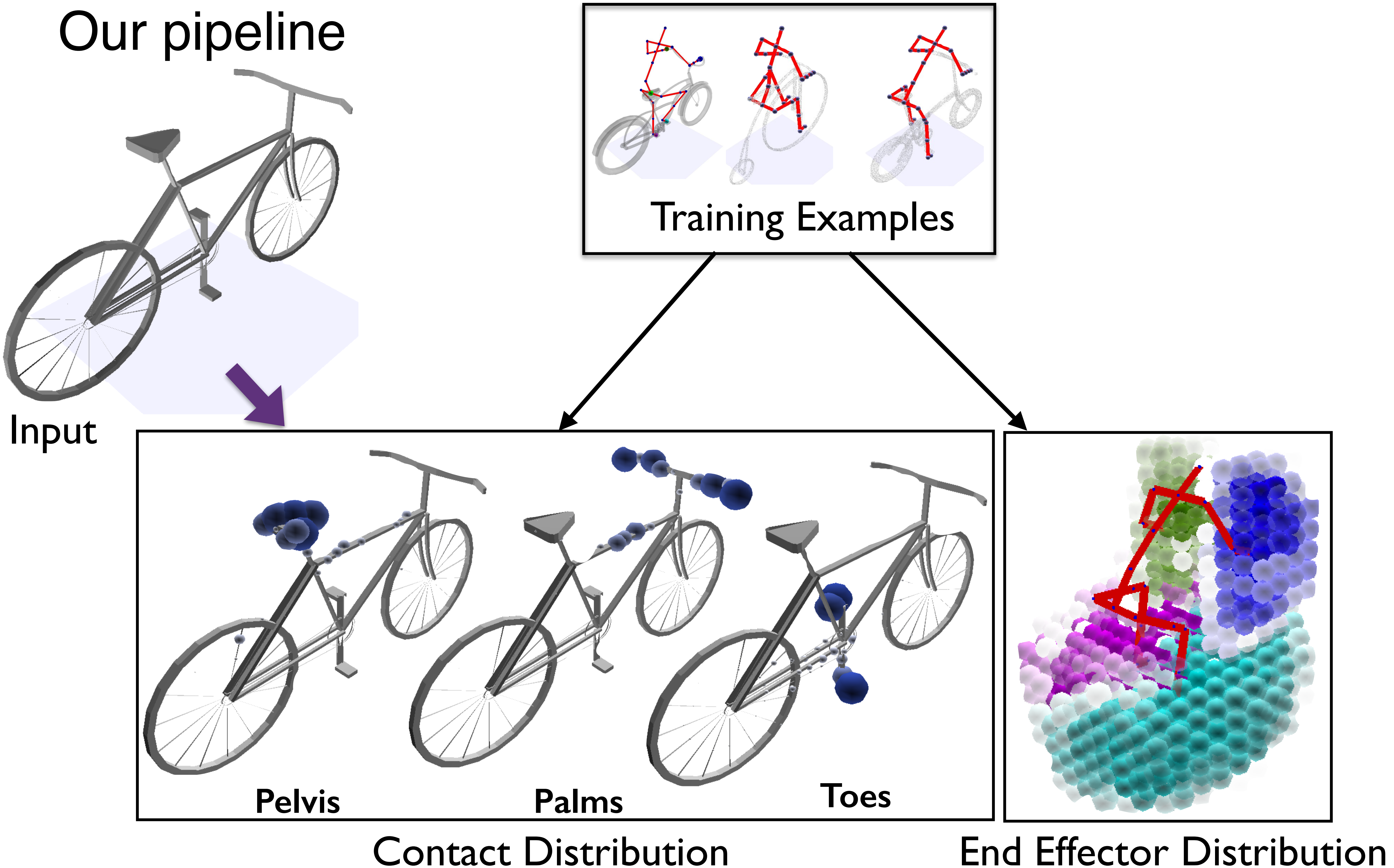
Pose Prediction

Our pipeline

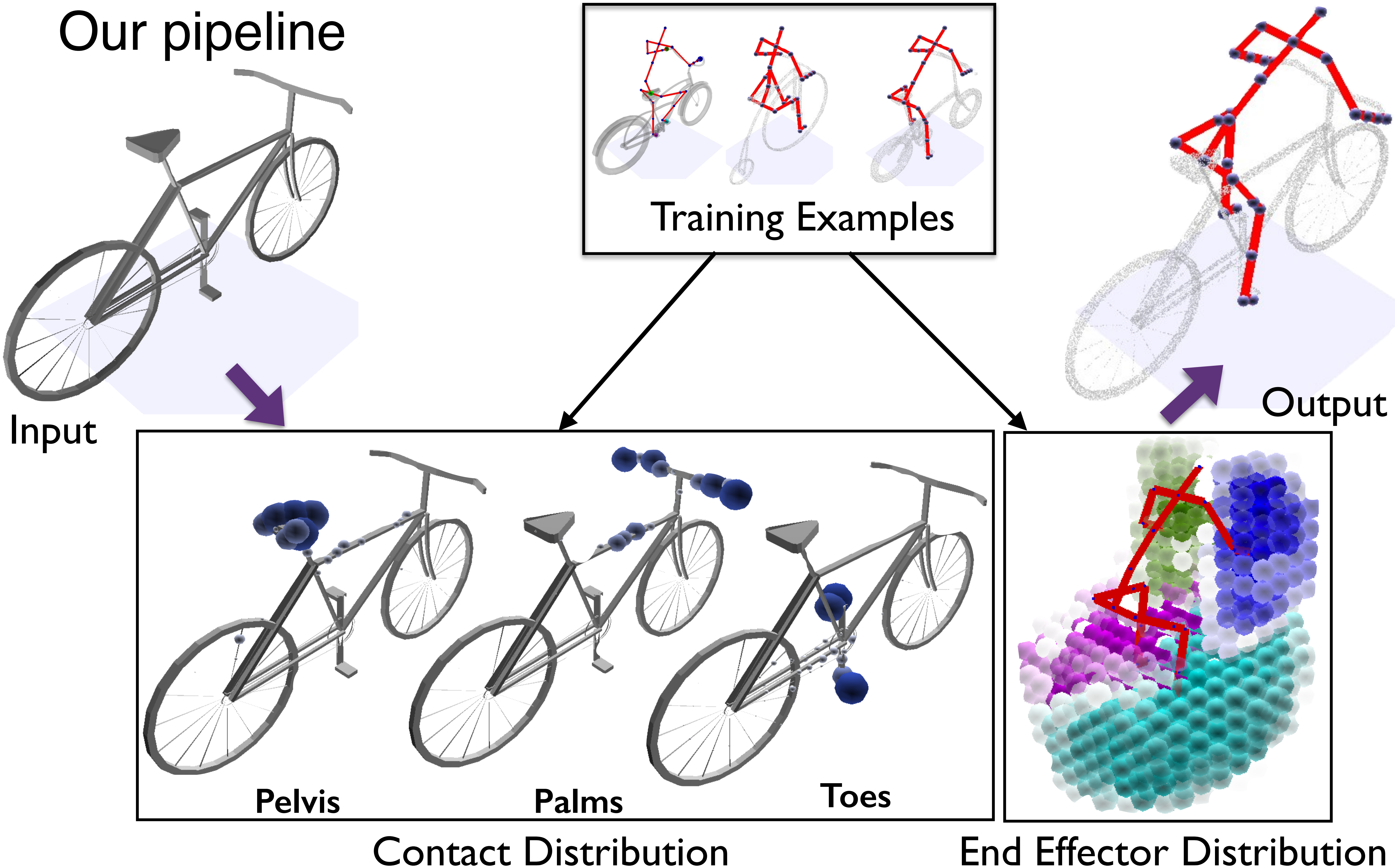


Pose Prediction

Our pipeline



Pose Prediction



Overview

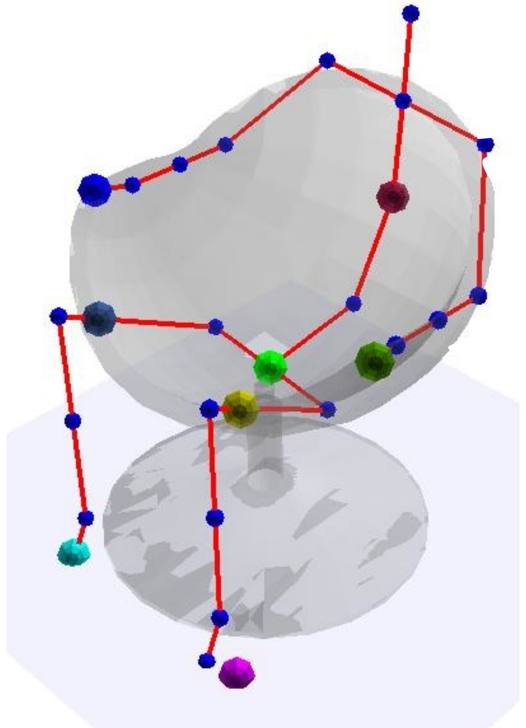
Introduction

Learning Affordance Model

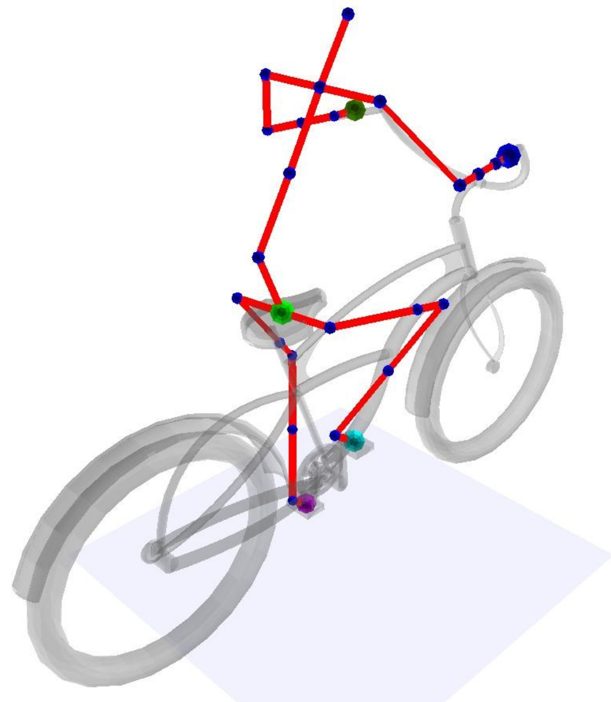
Pose Prediction

Results & Applications

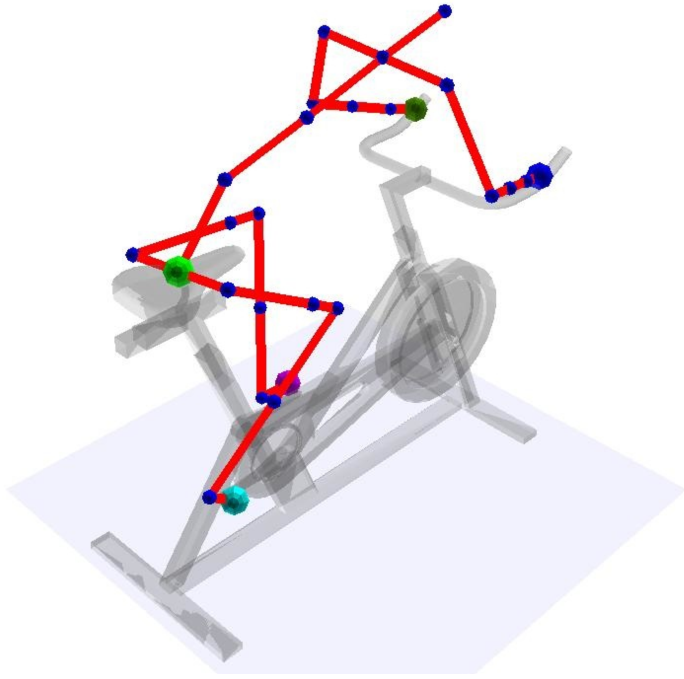
Datasets



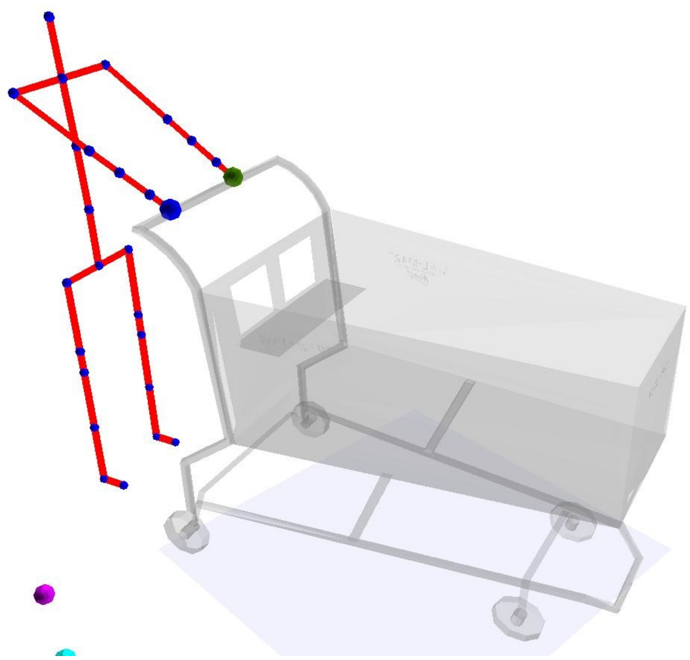
Chairs (30)



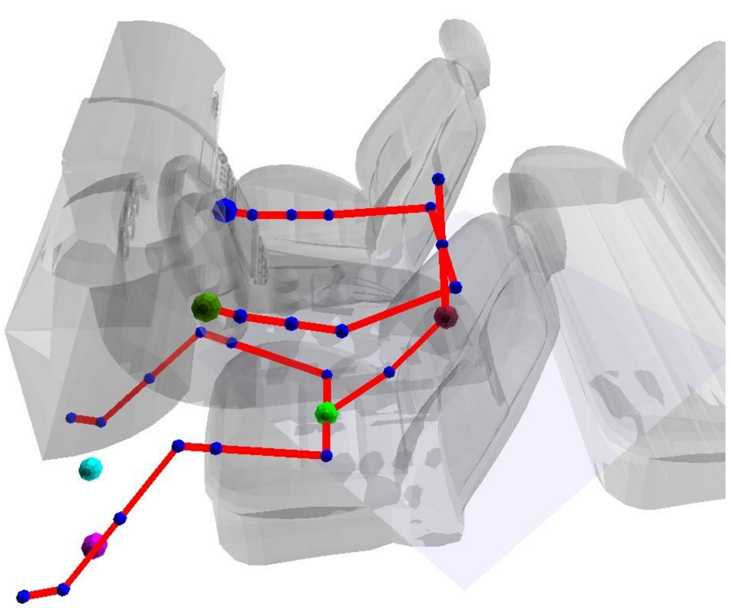
Bicycles (30)



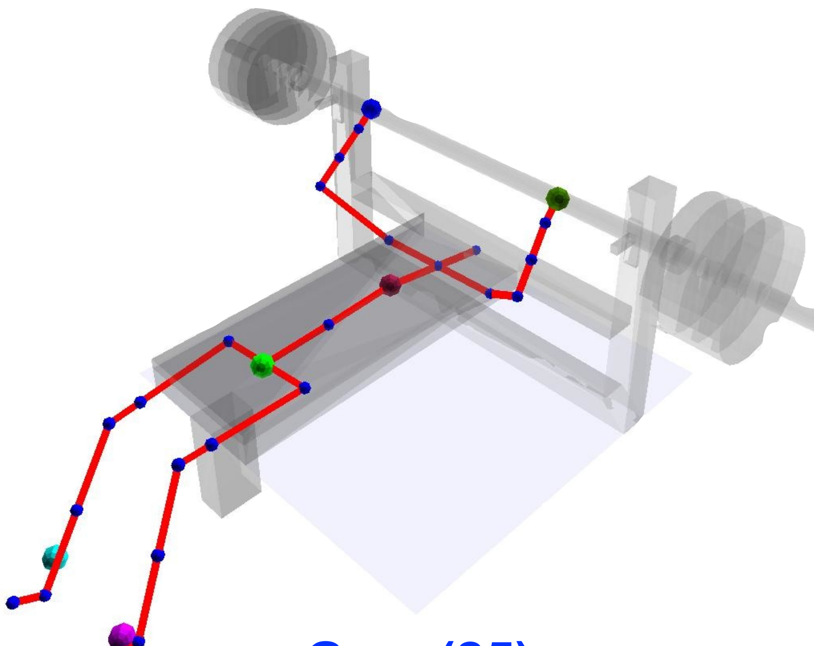
Bipedals (30)



Carts (11)

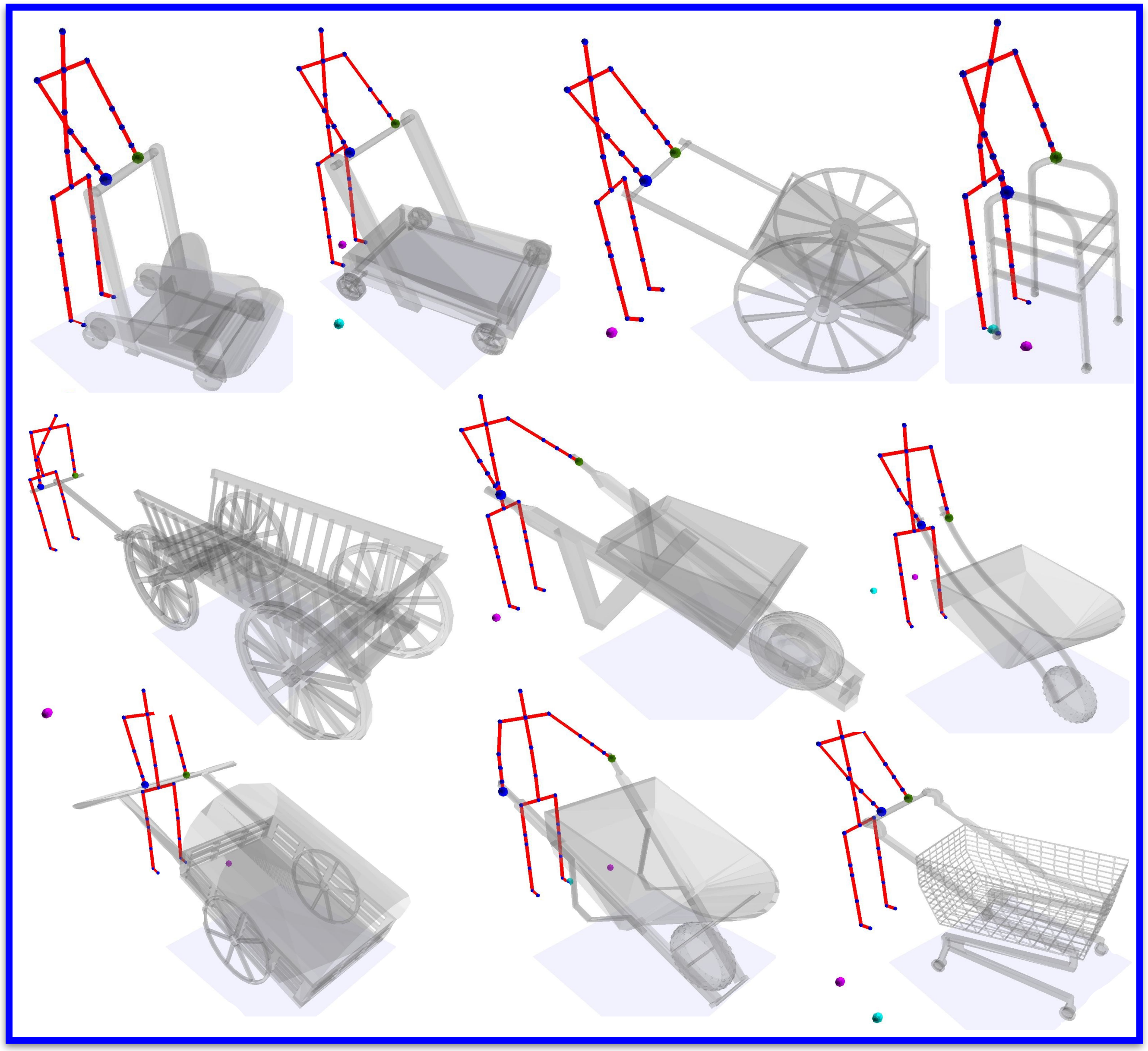


Cockpits (21)



Gym (25)

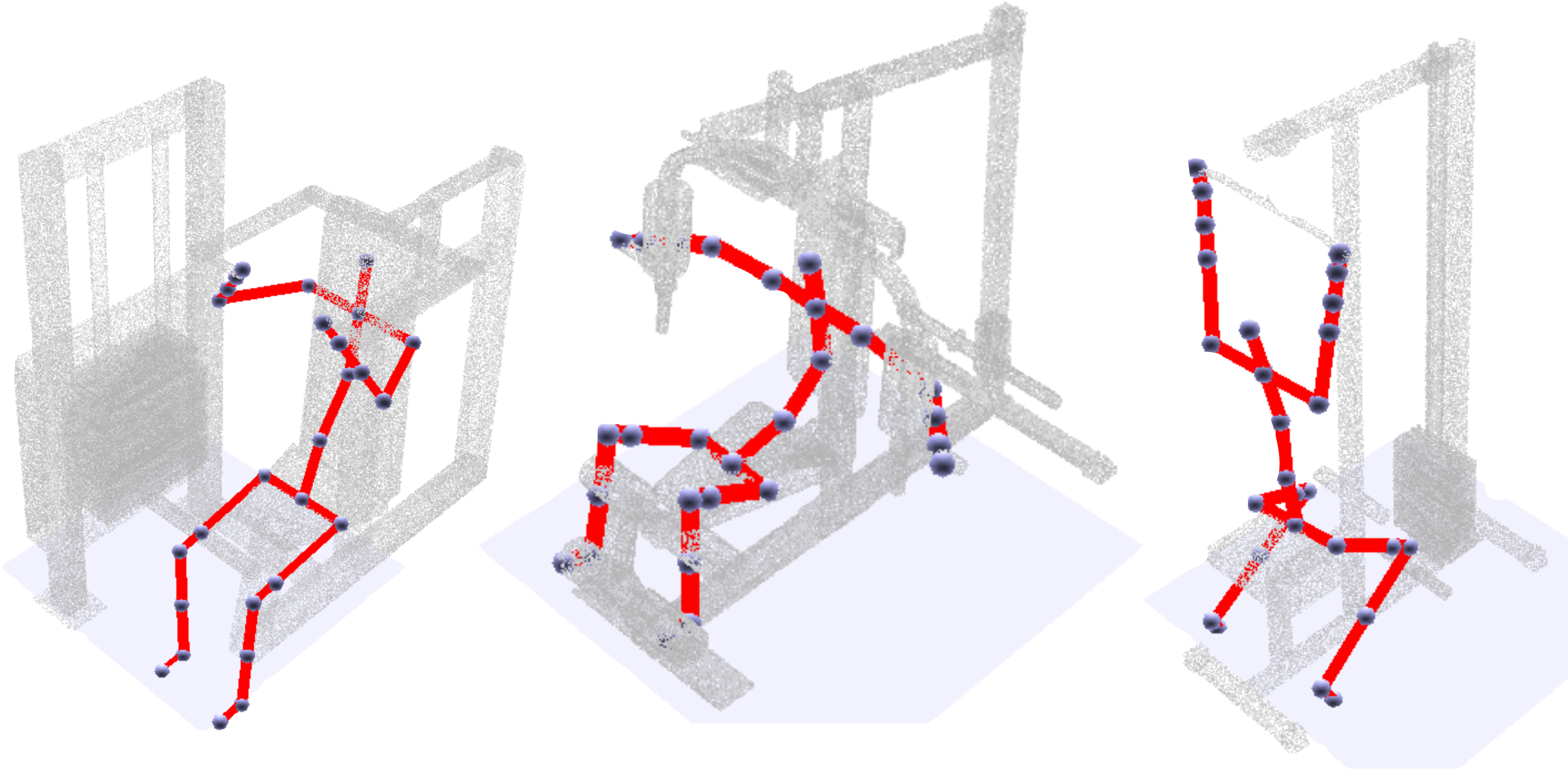
Leave-one-out Results



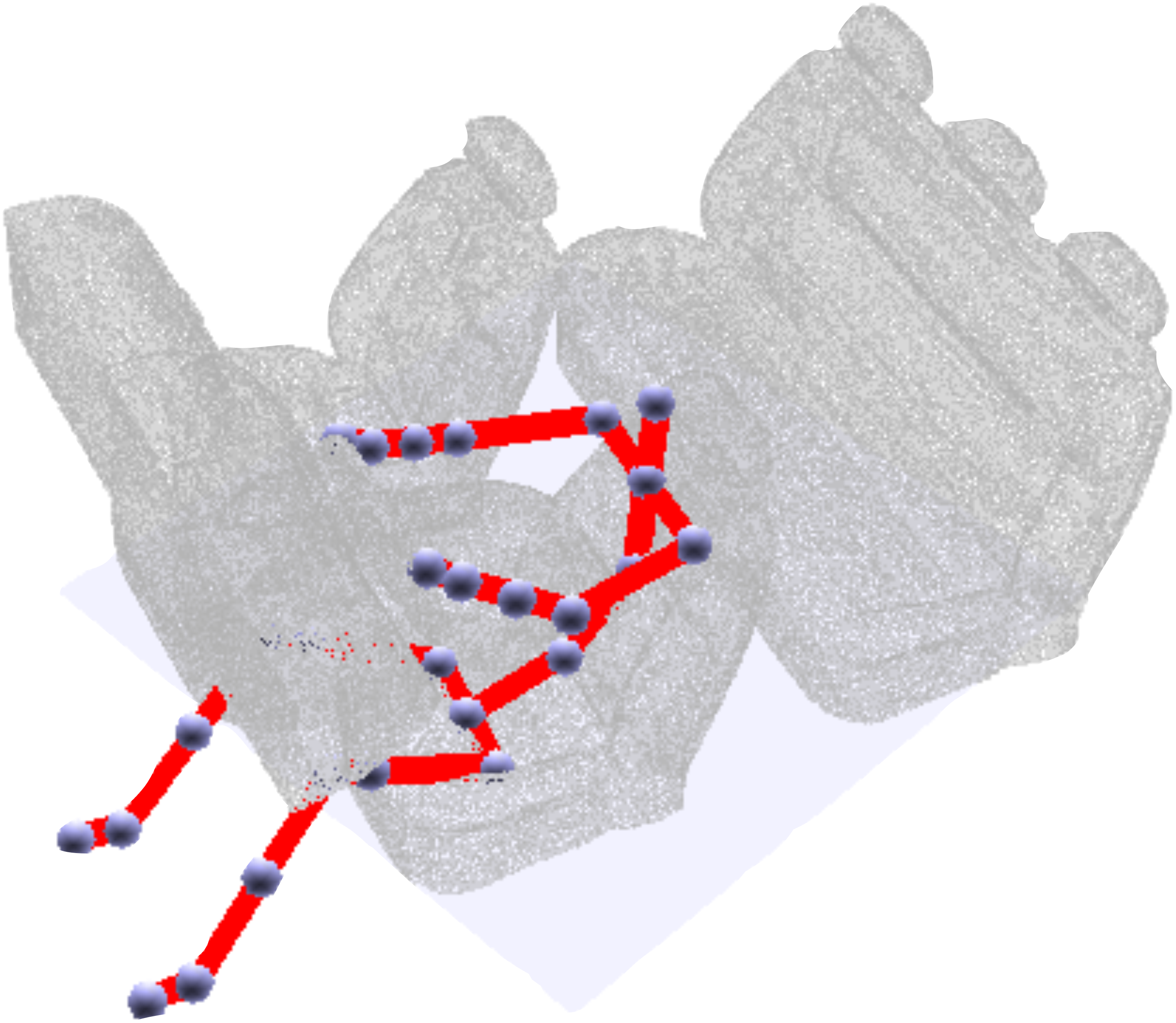
Training Data (10)

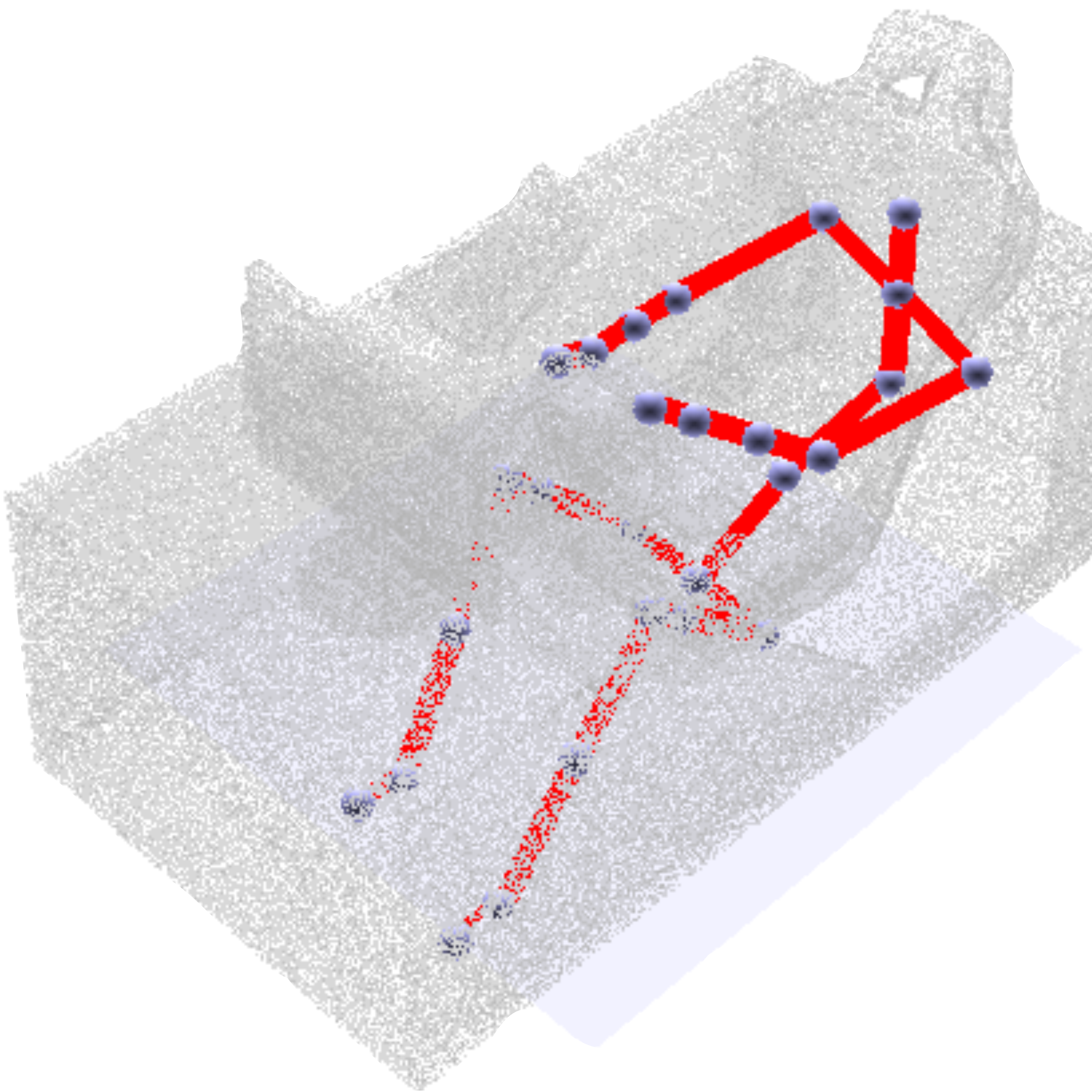
Test Data (1)

Pose Prediction



Pose Prediction



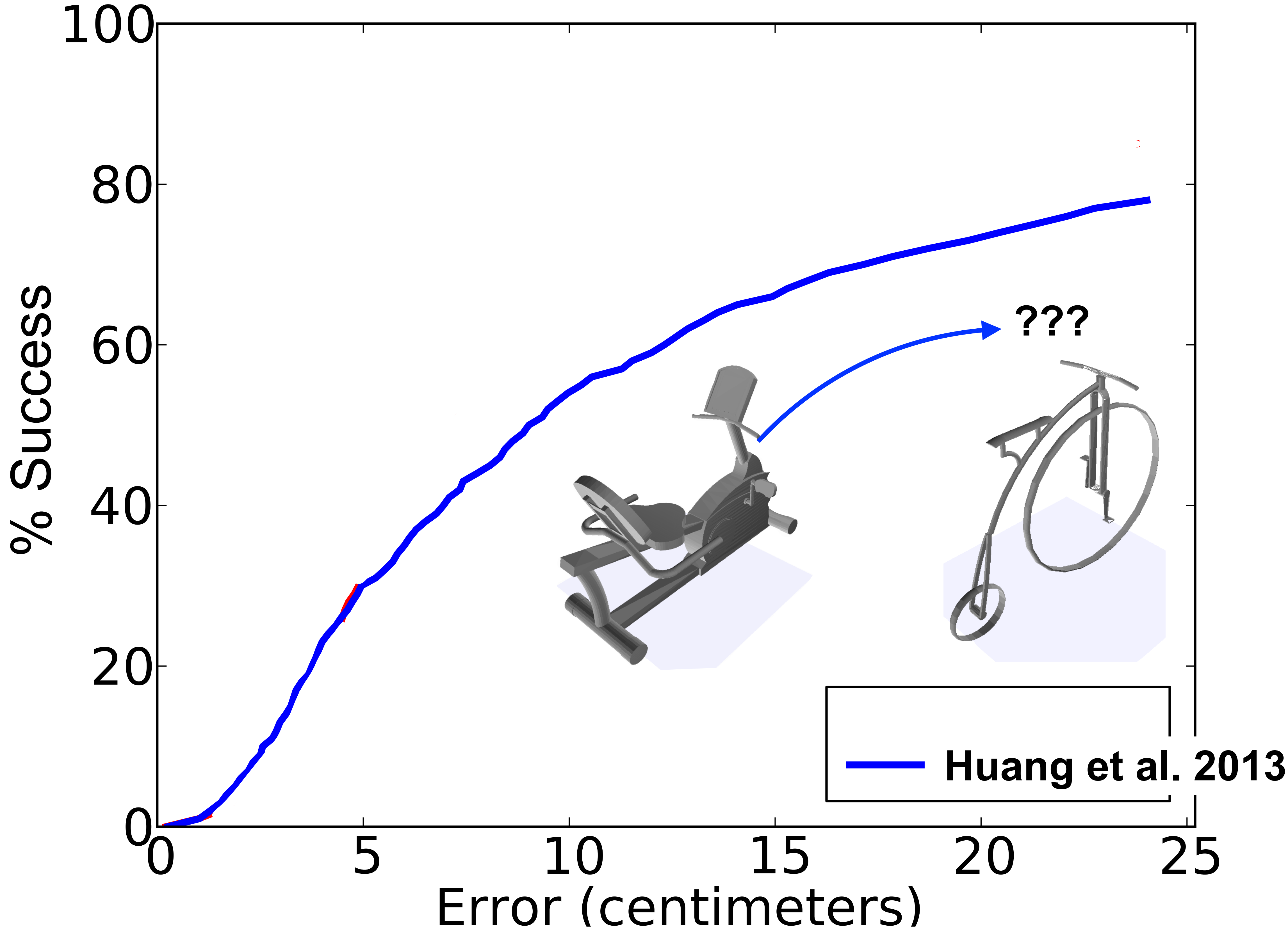


Applications

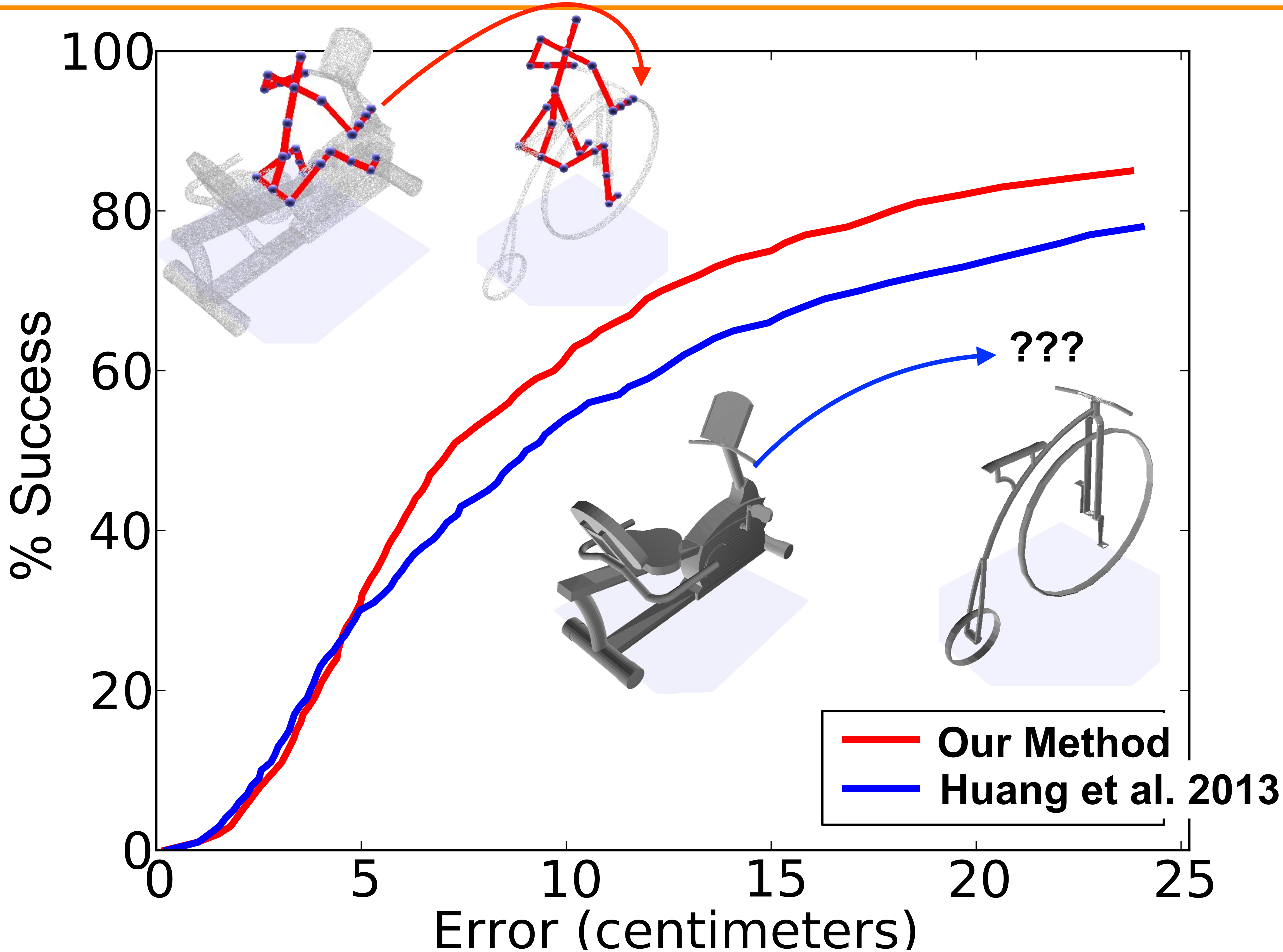
Examples

- Sparse Correspondence
- Salience Estimation
- Shape Retrieval

App: Sparse Correspondence



App: Sparse Correspondence



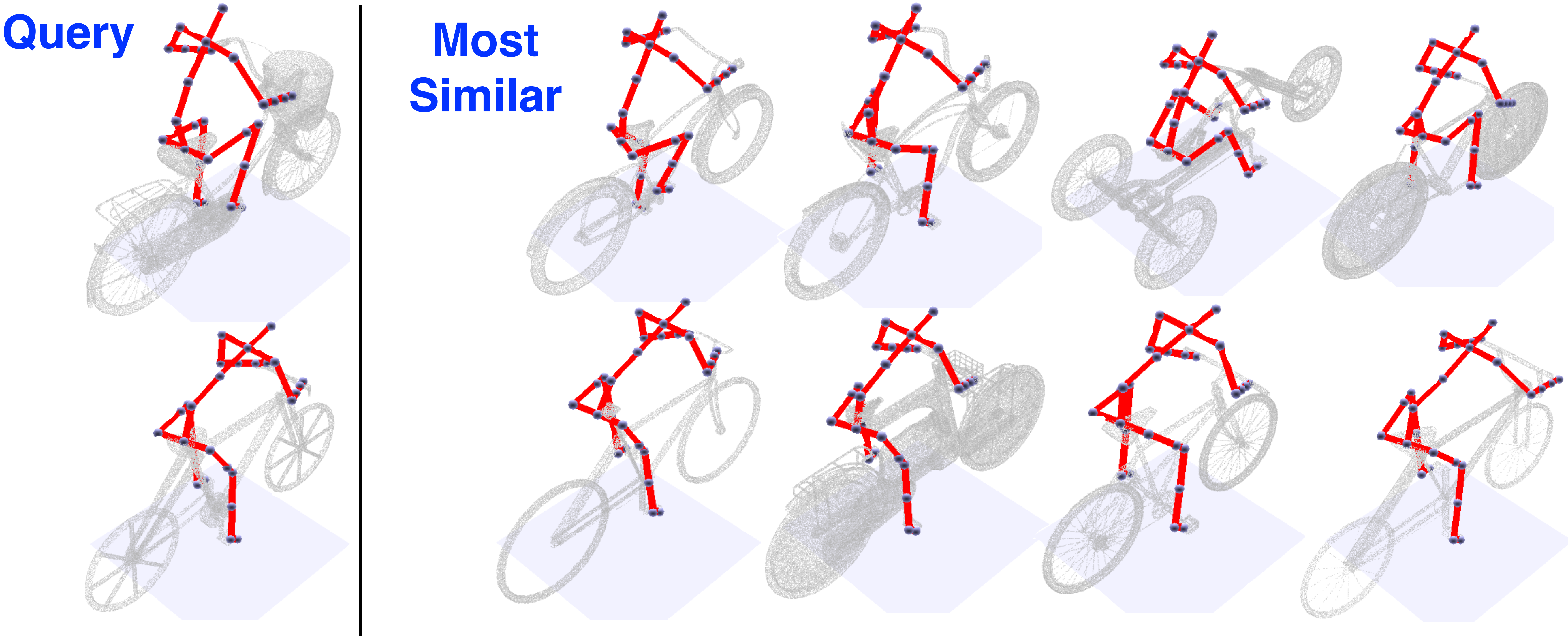
App: Salience Estimation

Mesh Saliency [Lee et al. 2005]

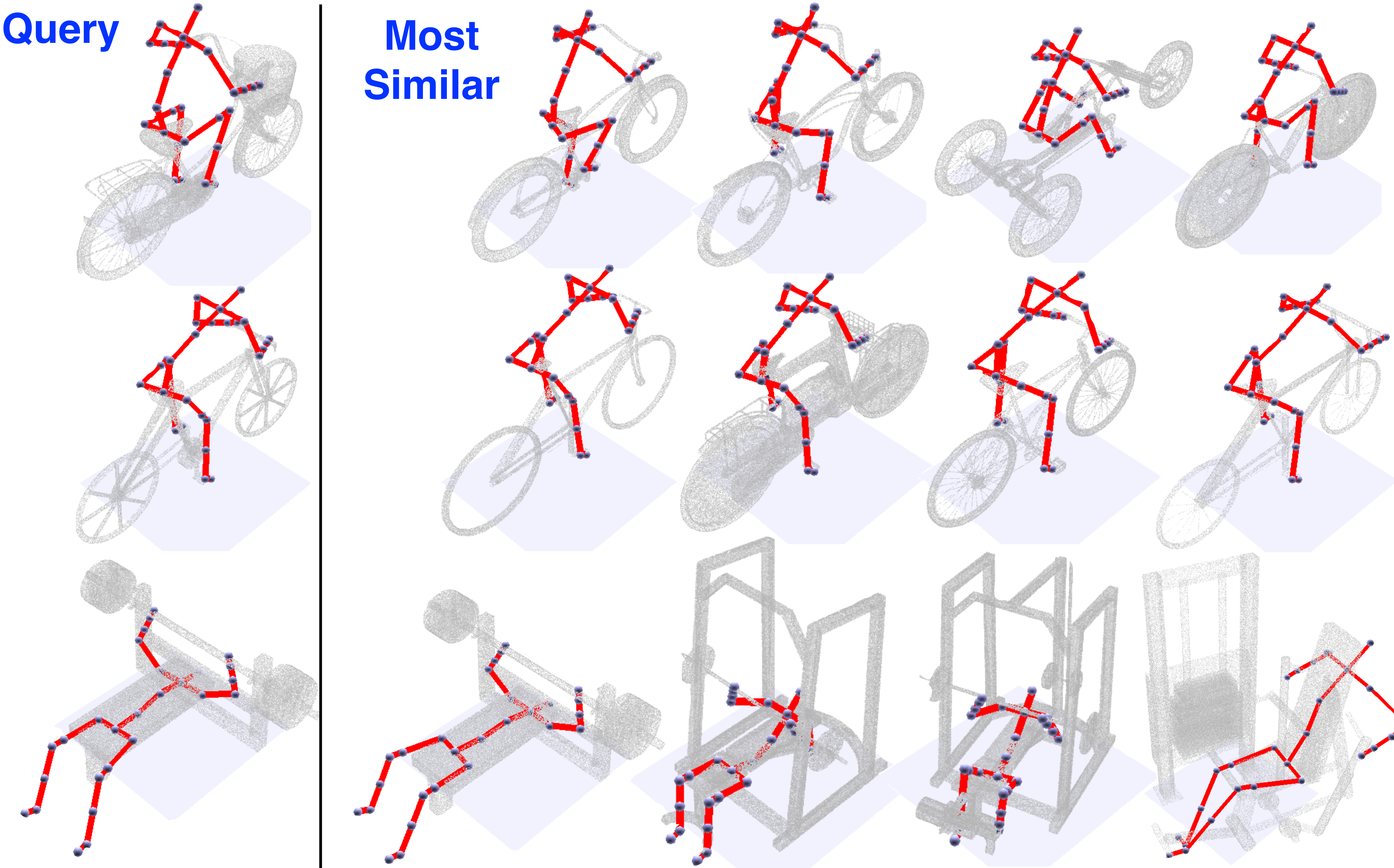
App: Salience Estimation

Human-centric Saliency [Our method]

App: Shape Classification & Retrieval



App: Shape Classification & Retrieval

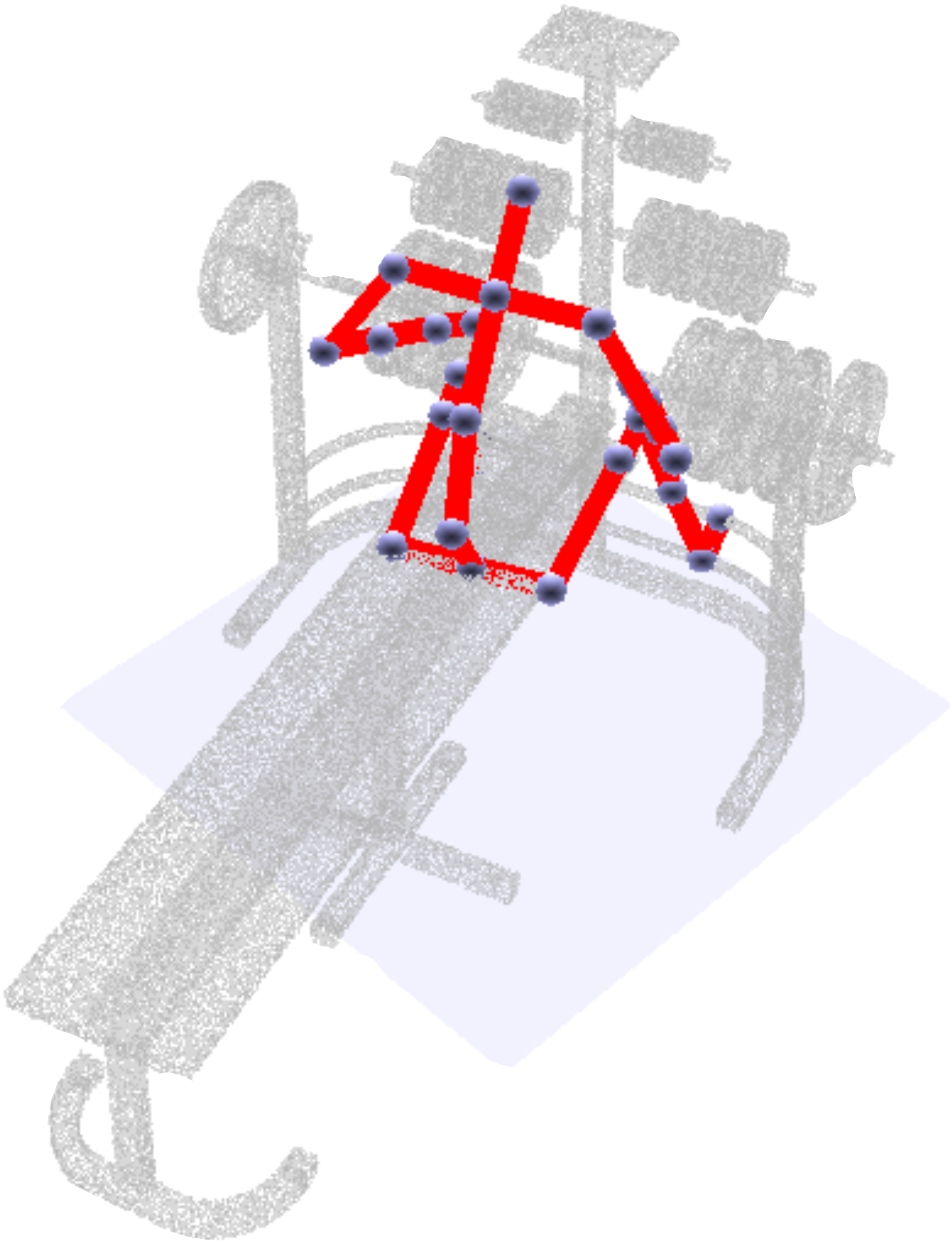


Summary

Human-Centric Shape Analysis

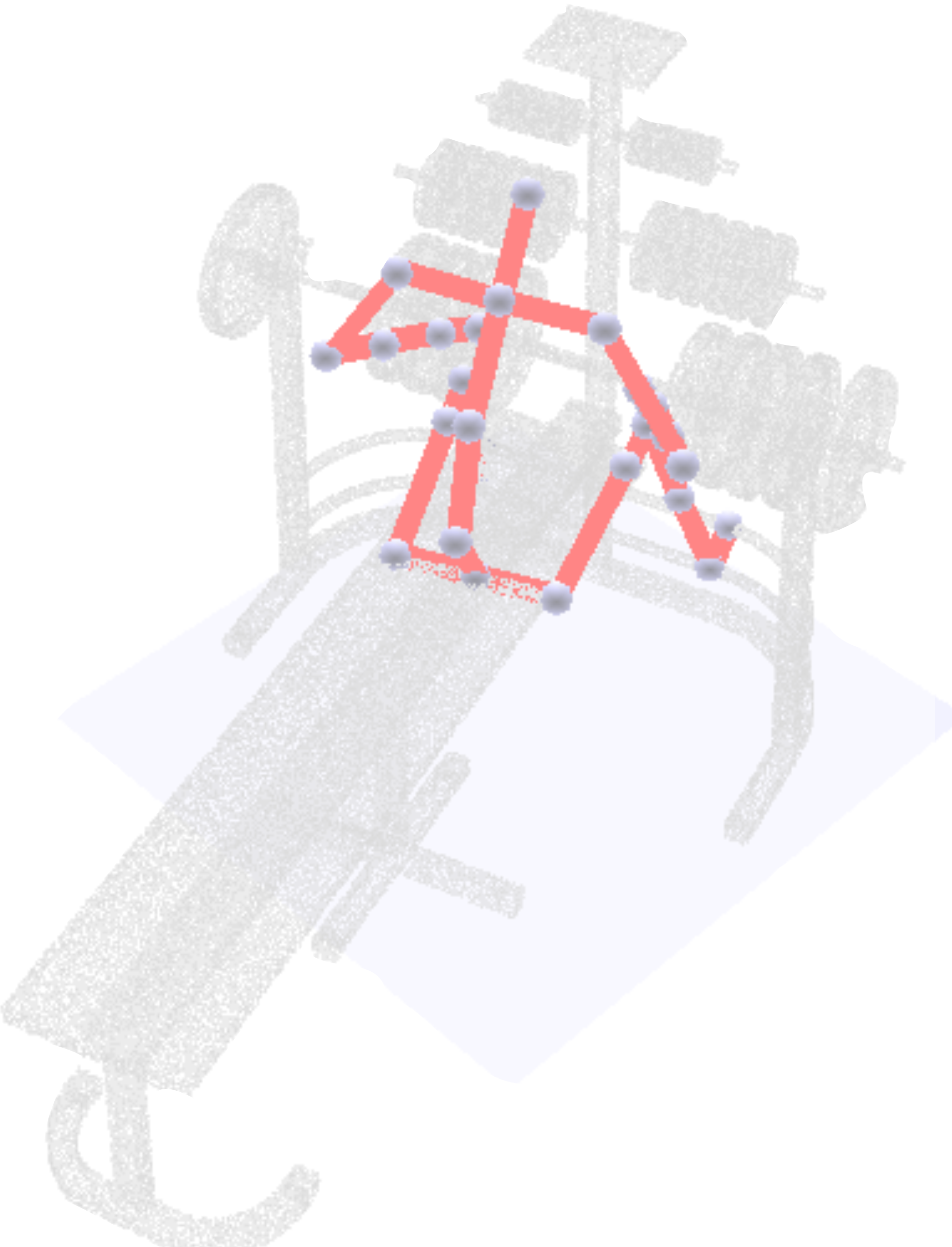
- Affordance is an intrinsic property of a shape
- Efficient optimization by pre-computing end effector distribution
- Applications: correspondence, saliency, retrieval, …

Limitations



Semantics

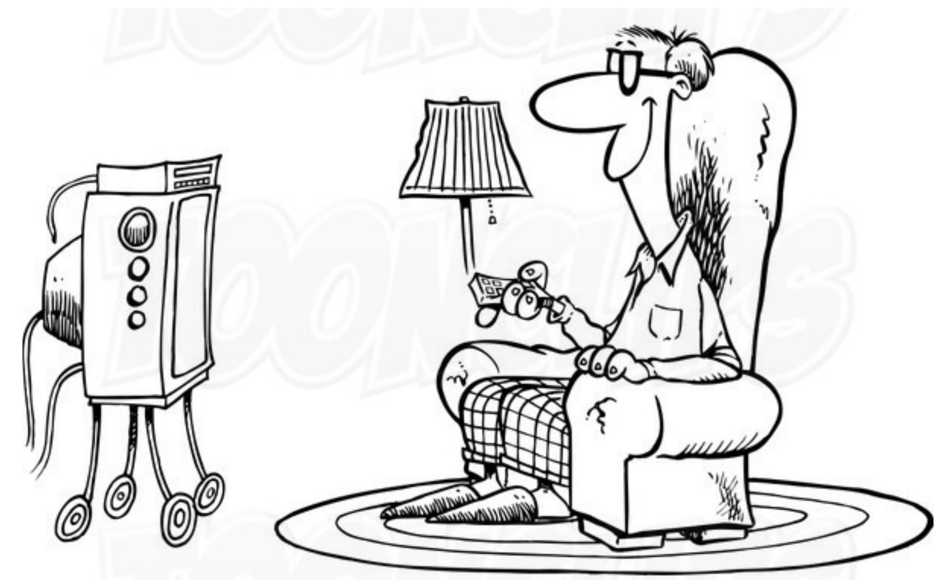
Limitations



Semantics

Future Work: predict a range of activities

Chair: sitting

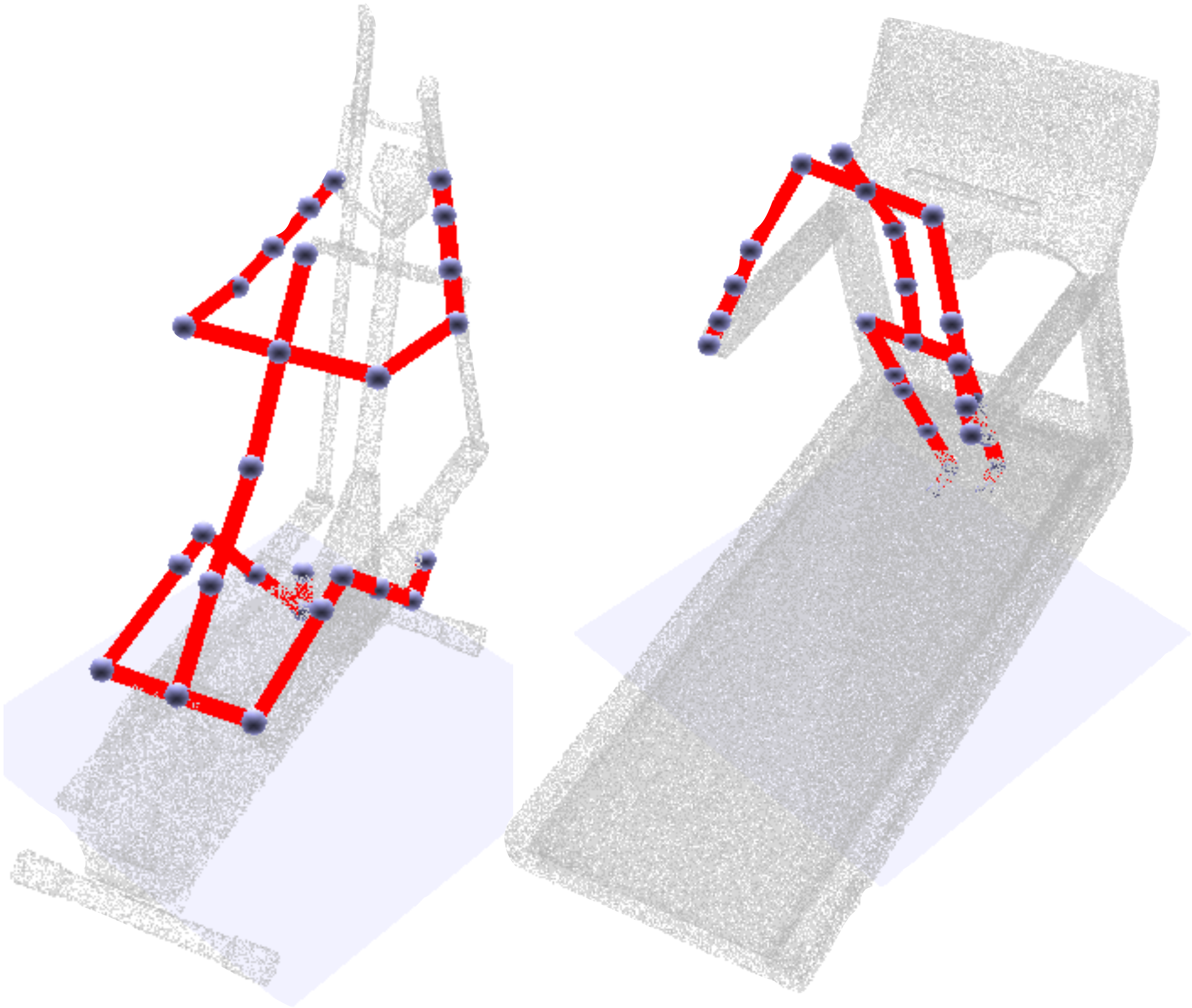


Chair: watching TV

Chair: pushing

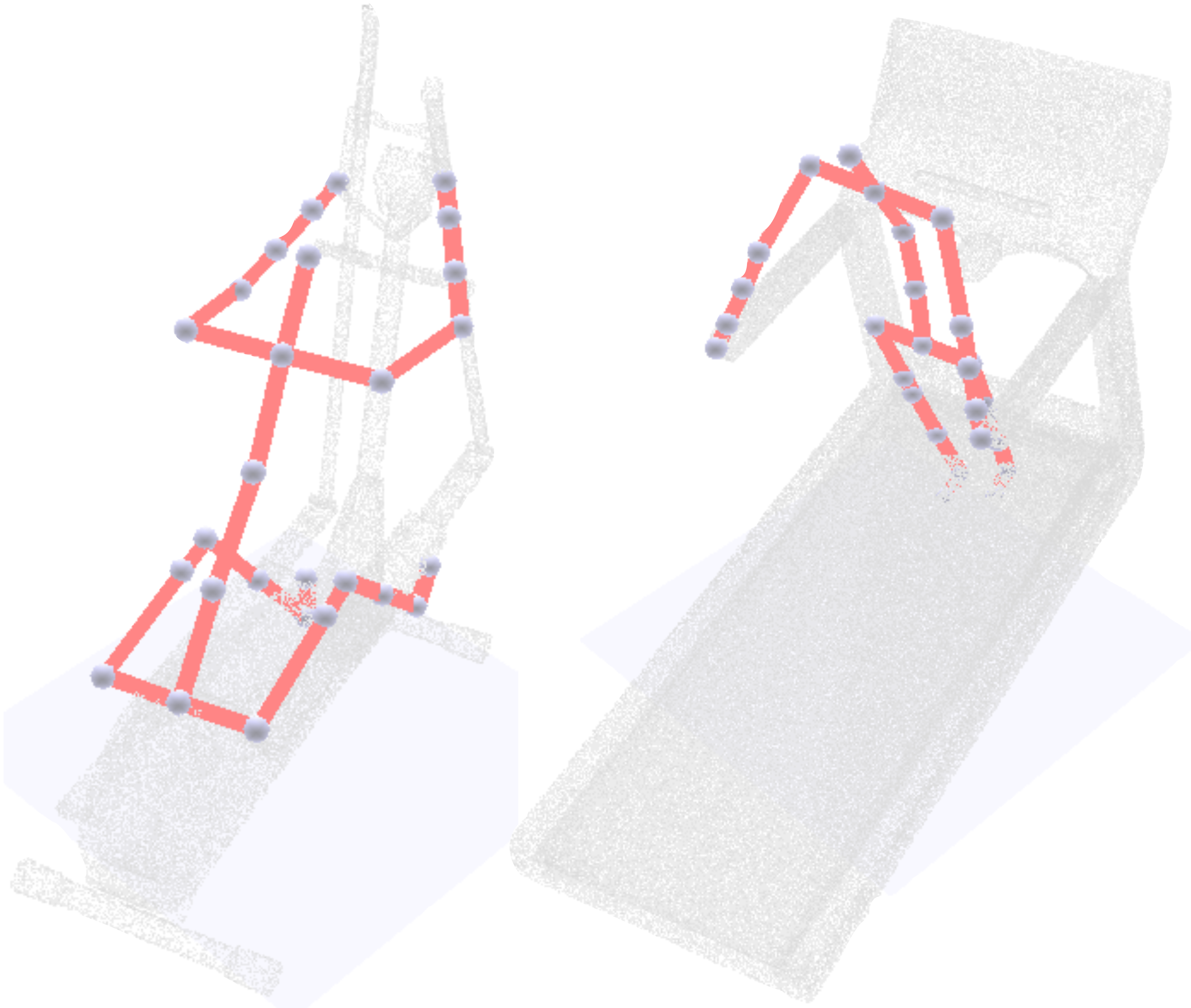
Chair: working

Limitations



Dynamics

Limitations



Dynamics

Future Work: model dynamic interactions

Future Applications

Functional Recognition

Populating Virtual Environments

Object Design

Acknowledgement

Code and discussions:

- Qixing Huang, Ashutosh Saxena, Peter Minary, Hao Zhang

Funding:

- NSF, Intel, Google, Adobe

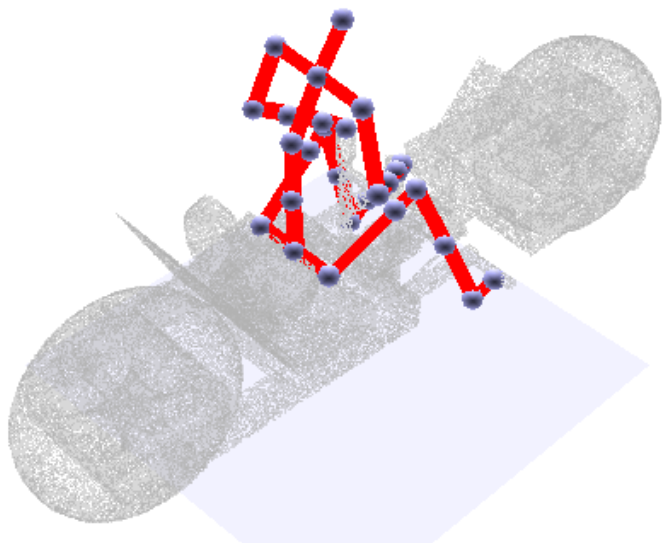
CODE AND DATA

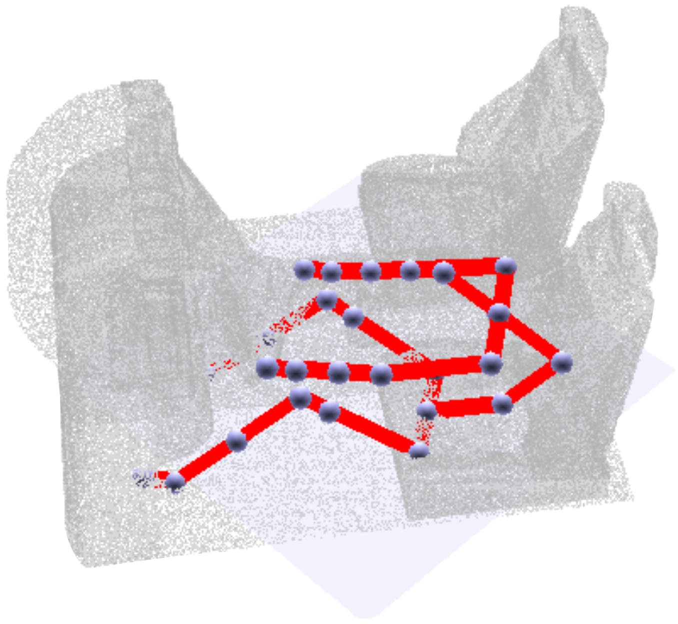
http://www.cs.princeton.edu/~vk/projects/Shape2Pose/

Thank You!

Comparison

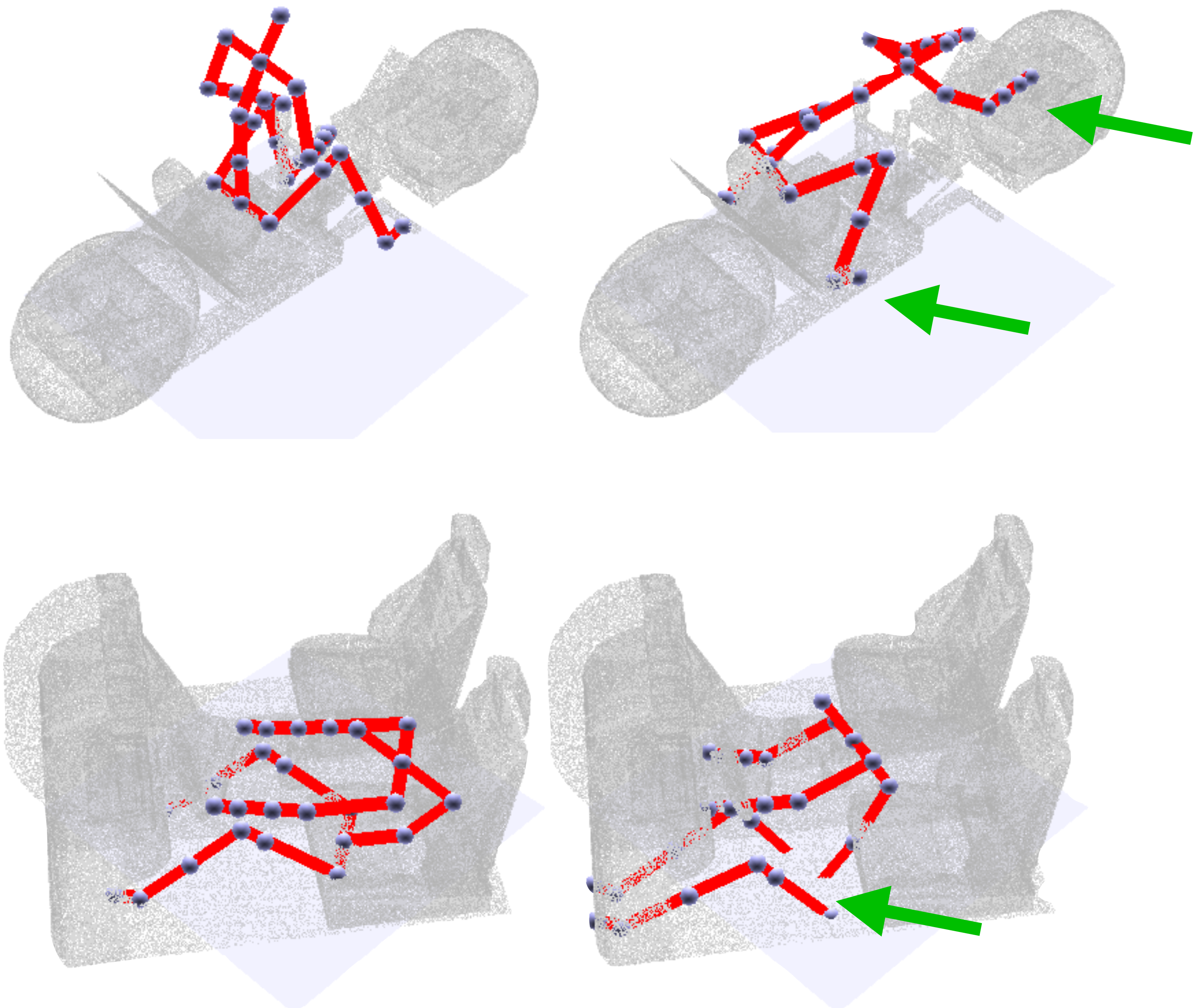
Our





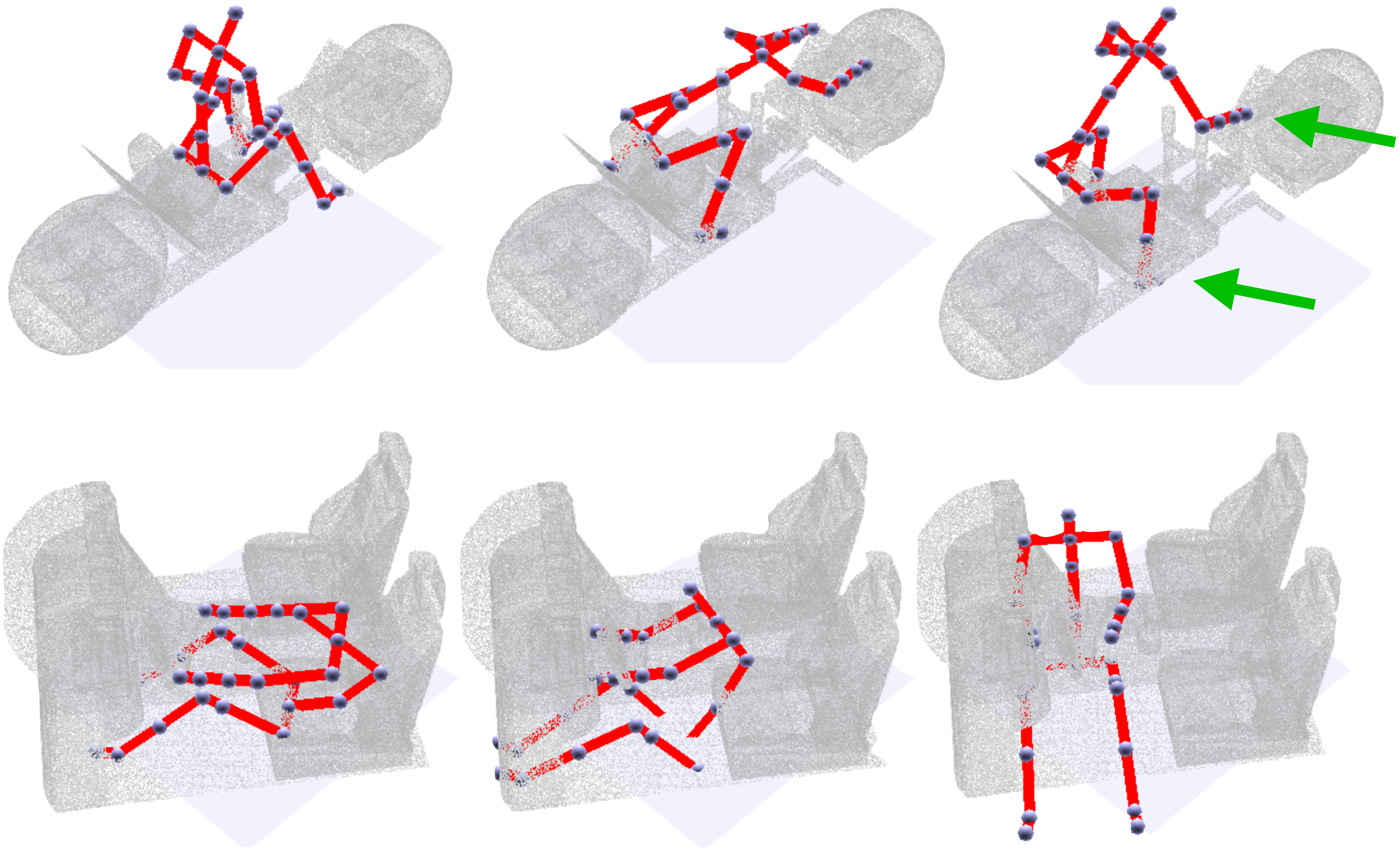
Comparison

Our — **Shape Matching**



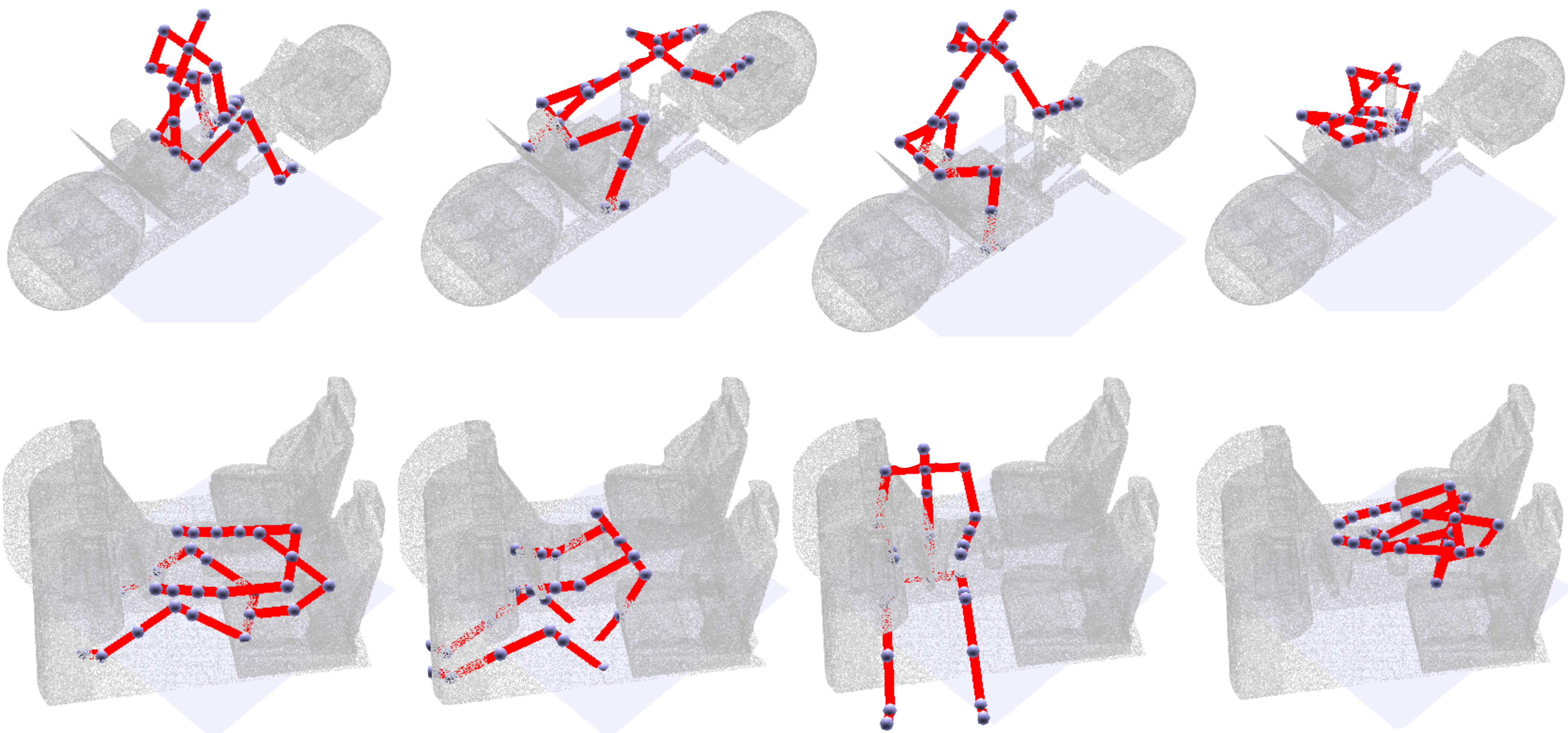
Comparison

Our | Shape Matching | Rigid Pose

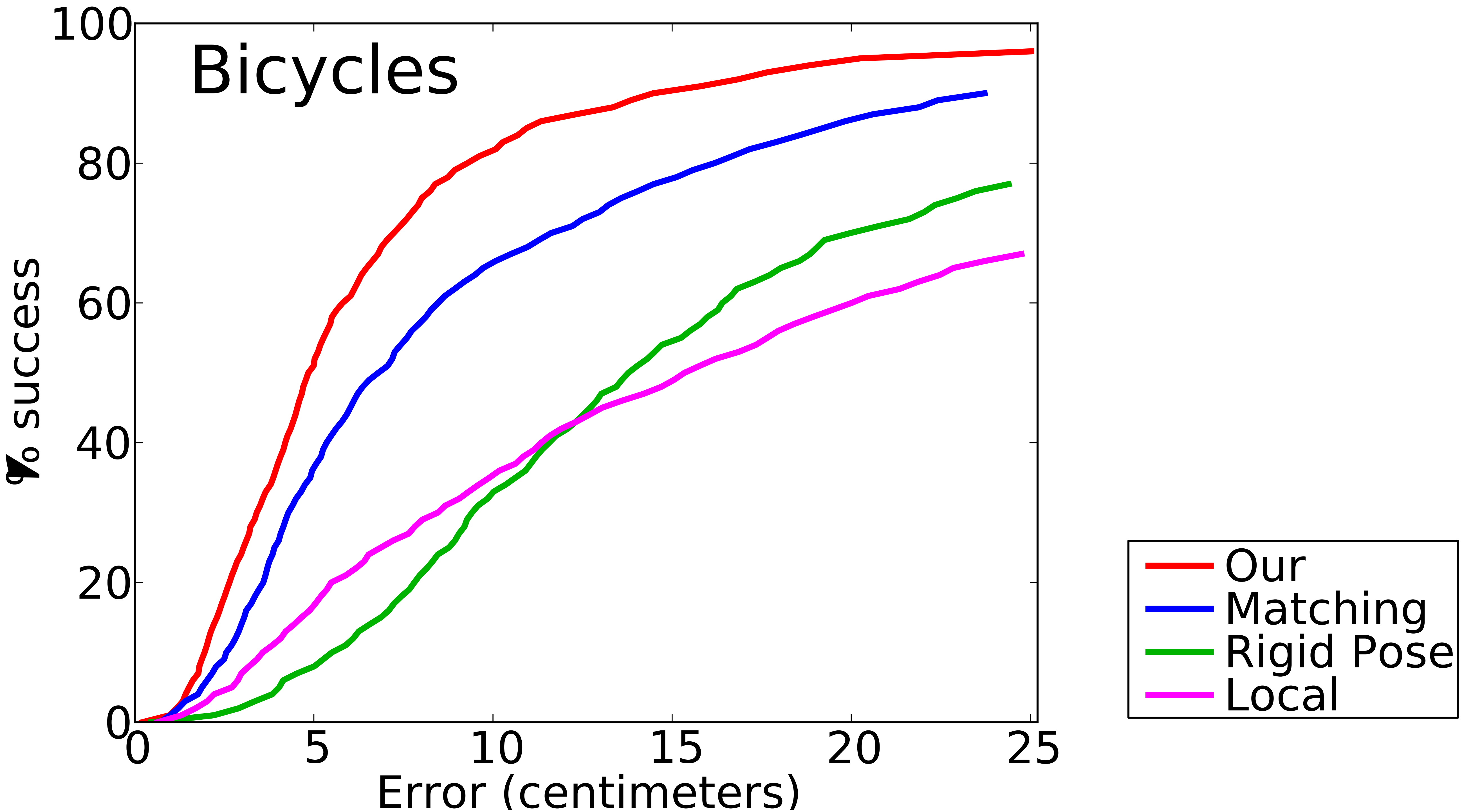


Comparison

Our | Shape Matching | Rigid Pose | Local

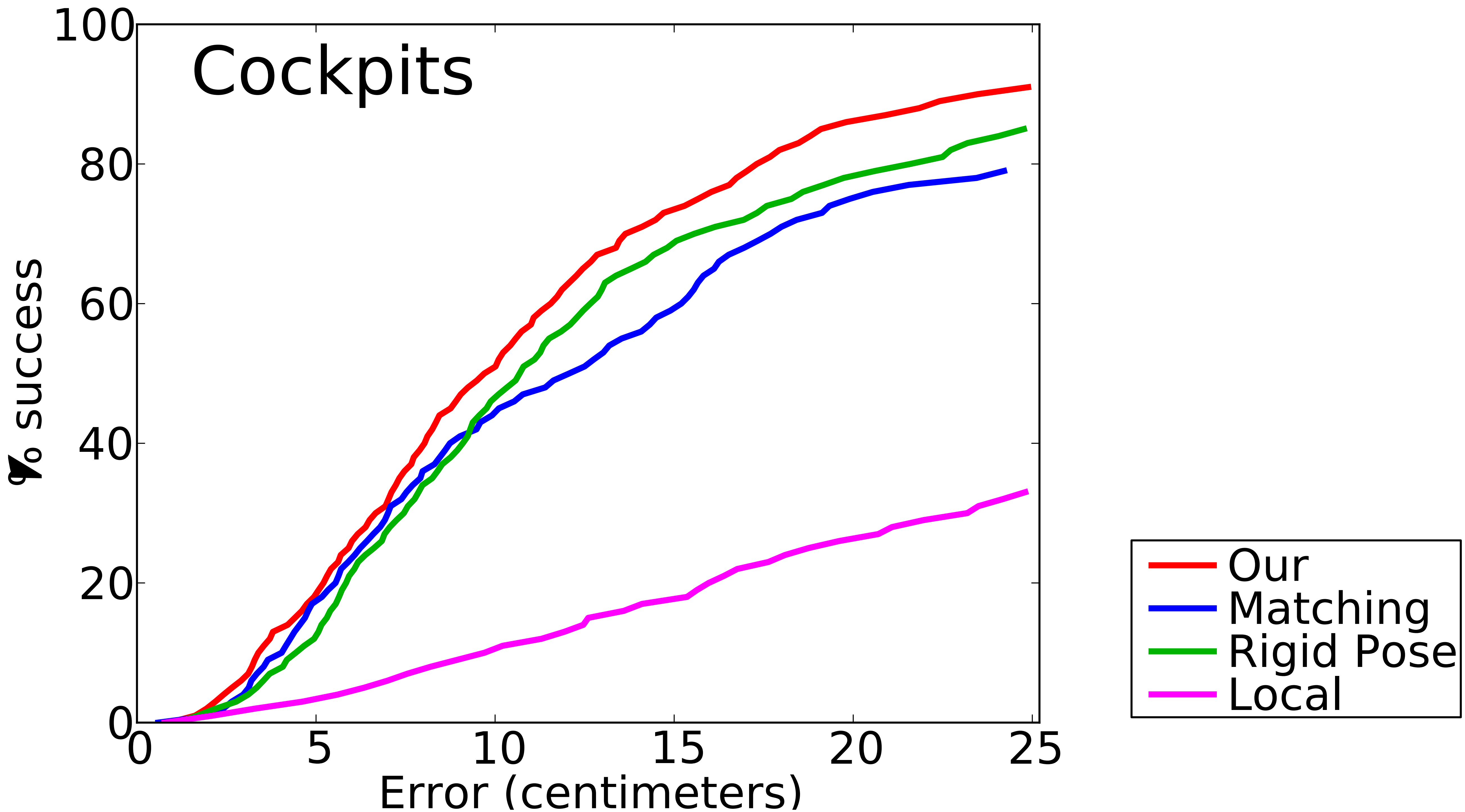


Comparison

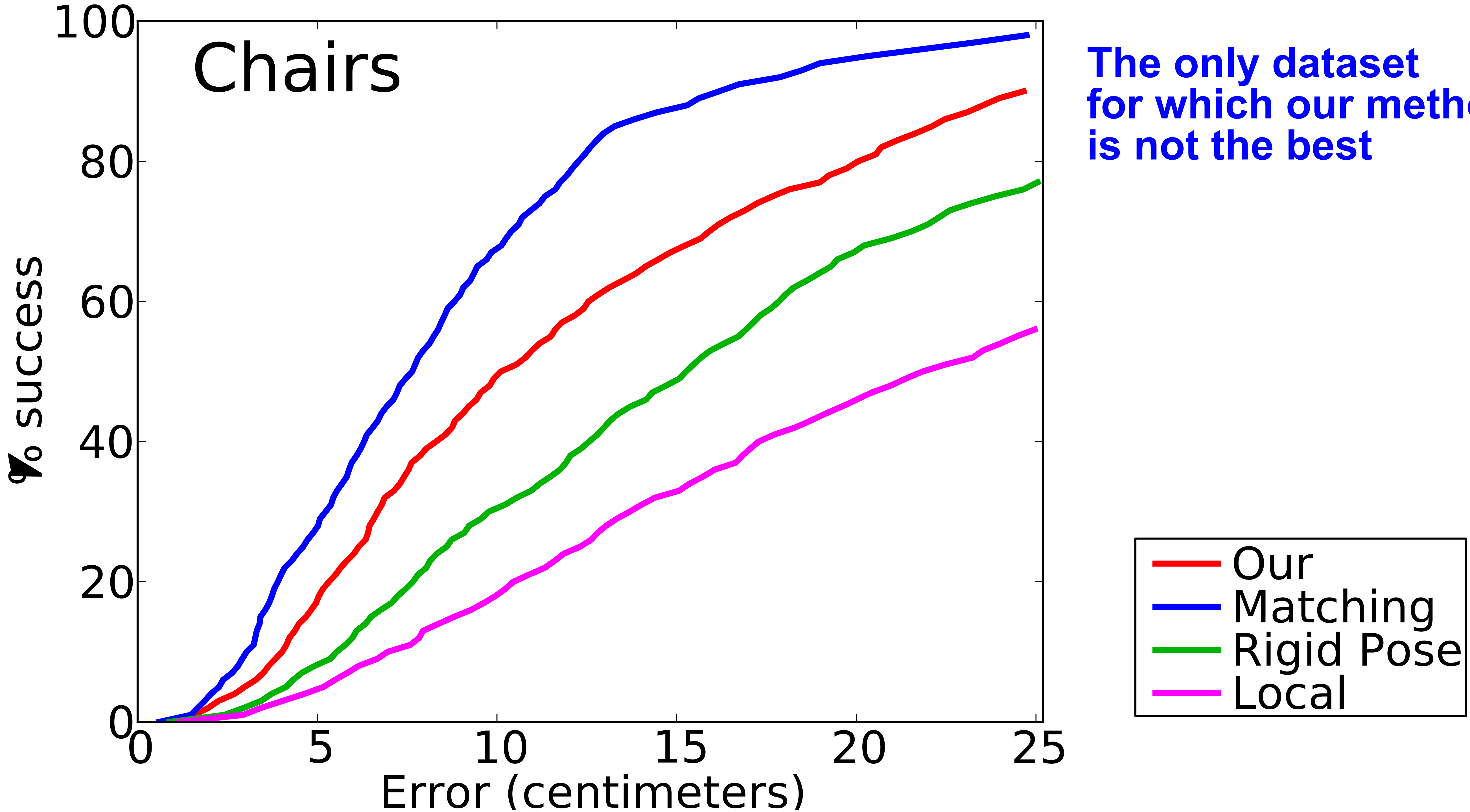


Bicycles
% success
Error (centimeters)
0
20
40
60
80
100
0
5
10
15
20
25
Our
Matching
Rigid Pose
Local

Comparison


Cockpits
100
80
60
40
20
0
0
5
10
15
20
25
% success
Error (centimeters)
Our
Matching
Rigid Pose
Local

Comparison



The only dataset for which our method is not the best

Timing and Complexity

Timing

Data	N	Prep	Train	Opt
Bicycles	30	80s	115s	130s
Bipedals	30	225s	200s	590s
Cockpits	21	1150s	550s	970s
Carts	11	235s	25s	15s
Chairs	30	50s	60s	80s
Gym Equipment	25	345s	270s	500s

Complexity

1. For each candidate contact alignment → N_{cand}
2. For each rotation around "up" → $N_{rot}=32$
3. Greedily add best-energy points → N_{cand}
4. Sort poses by lower bound energy
5. Run IK - find exact poses → At most $N_{cand} \cdot N_{rot}$
   Compute full energy
6. If Full Energy < Lower Bound - terminate

$O(N_{cand}^2)$