Approaches to Probabilistic Model Learning for Mobile Manipulation Robots

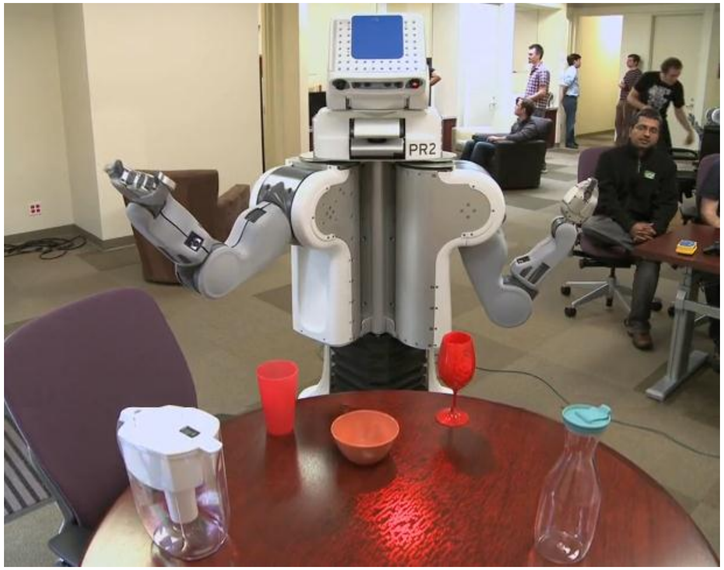
Jürgen Sturm
University of Freiburg
(now at Technical University of Munich)

PhD Supervisor: Wolfram Burgard

Motivation

What could flexible service robots do for us?

- Fetching and carrying things
- Tidying up
- Cleaning



Motivation

At home

In healthcare

In SMEs

To accomplish these tasks, service robots need the capability to interact with cabinet doors and drawers.

Question:
How to model such articulated objects?

Motivation

Goal:

Enable service robots to operate articulated objects.

Problem:

The work space of the robot is unknown at design time.

Challenge:

Robot needs to learn the required models on site.

[Sturm et al., IJCAI'09]

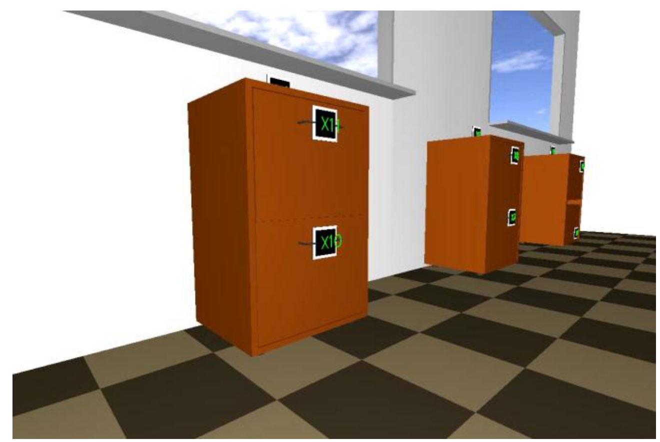
Problem Definition

- Given a sequence of pose observations of an articulated link

$$\mathcal{D} = \mathbf{z}_1, \ldots, \mathbf{z}_n \text{ with } \mathbf{z}_i \in \mathbb{SE}(3)$$

- Estimate the kinematic model

$$\hat{\mathcal{M}}, \hat{\theta} = \arg\max_{\mathcal{M},\theta} p(\mathcal{M}, \theta \mid \mathcal{D})$$



Bayesian Model Inference

Goal: Estimate

$$\hat{\mathcal{M}}, \hat{\theta} = \arg\max_{\mathcal{M},\theta} p(\mathcal{M}, \theta \mid \mathcal{D})$$

Split this using Bayesian inference into

- Step 1: Model Fitting

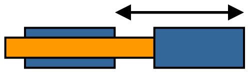
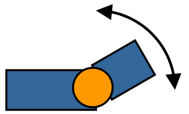
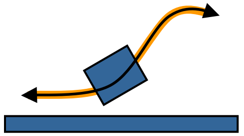
$$\hat{\theta} = \arg\max_{\theta} p(\theta \mid \mathcal{D}, \mathcal{M})$$

- Step 2: Model Selection

$$\hat{\mathcal{M}} = \arg\max_{\mathcal{M}} p(\mathcal{M} \mid \mathcal{D})$$

Step 1: Model Fitting

- Different objects require different models

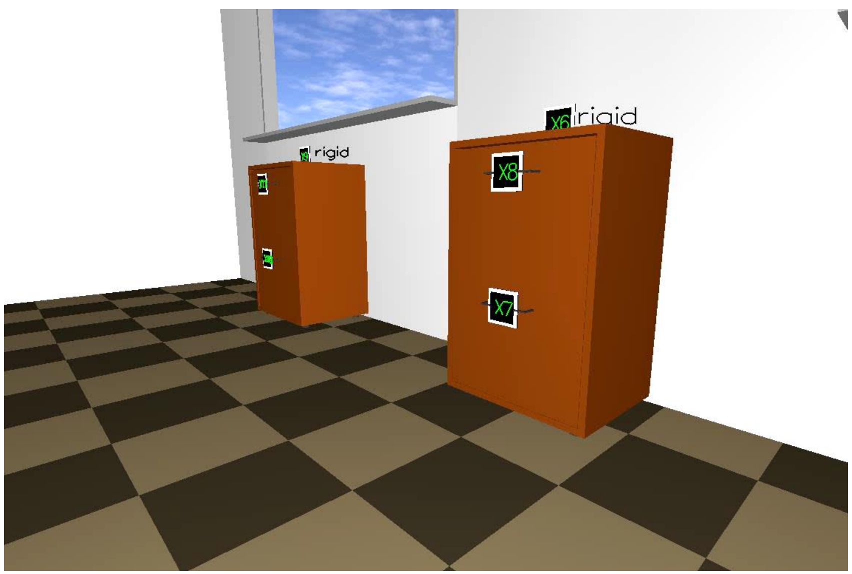
- Our set of candidate models
  - Rigid model 
  - Prismatic model 
  - Revolute model 
  - Gaussian process model 

[Sturm et al., IJCAI'09]

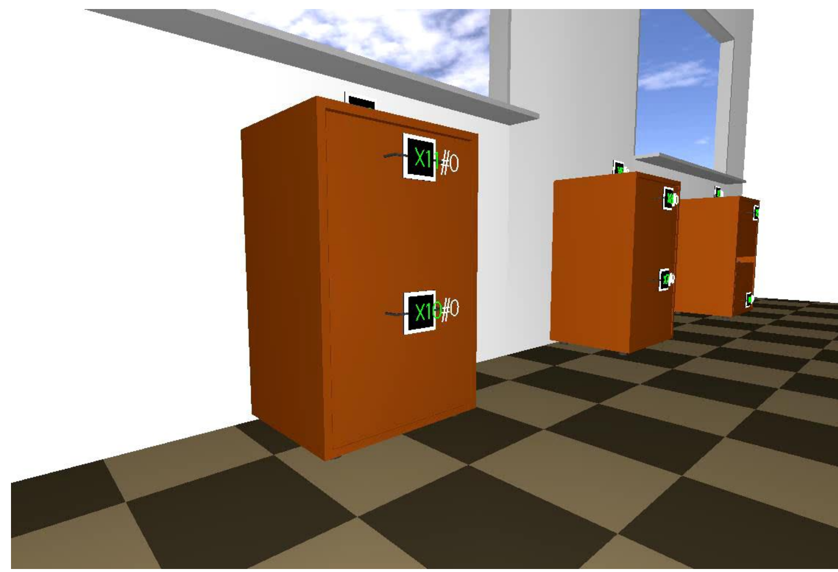
Parametric Models

- Noisy, outlier-corrupted data
- Robust estimation (MLESAC)
- Models are generative

Prismatic model

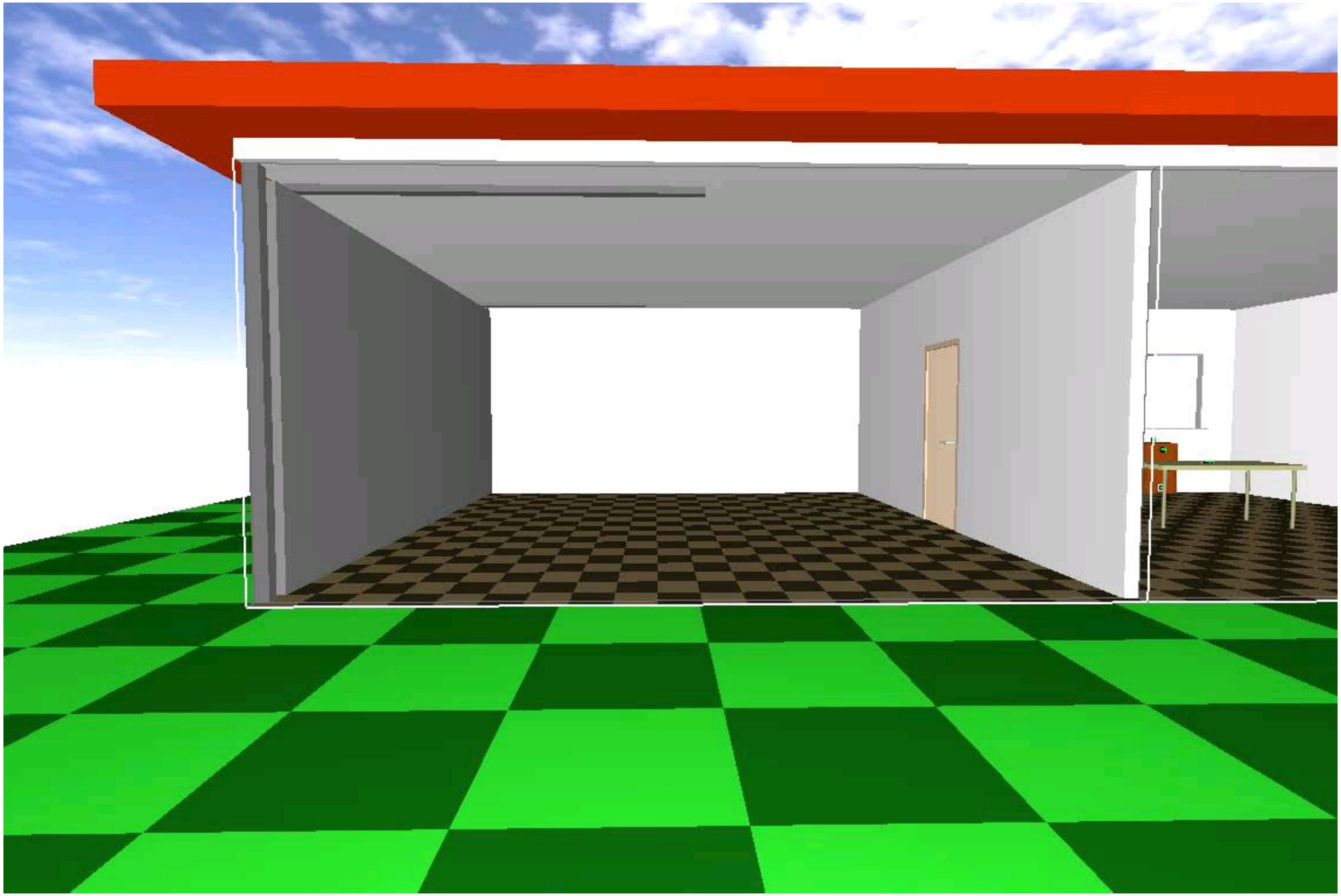


Revolute model



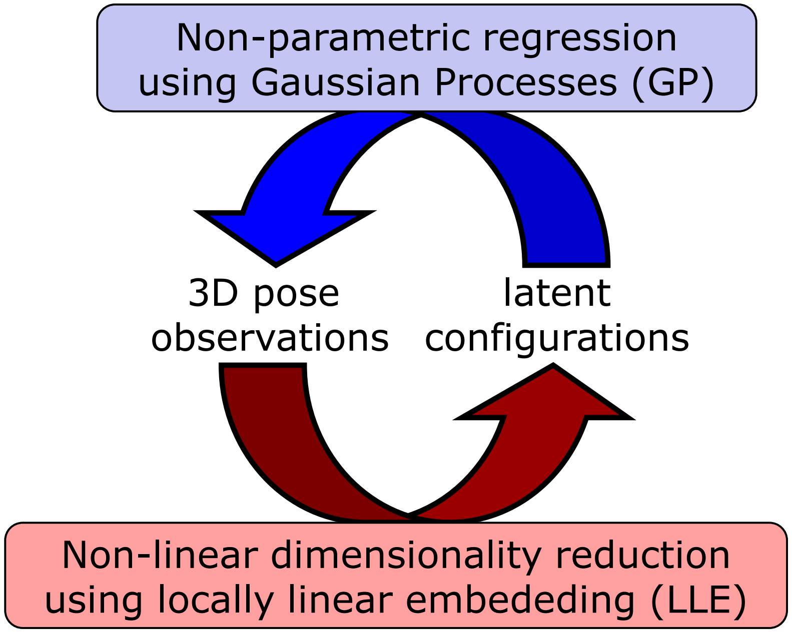
[Sturm et al., IJCAI'09]

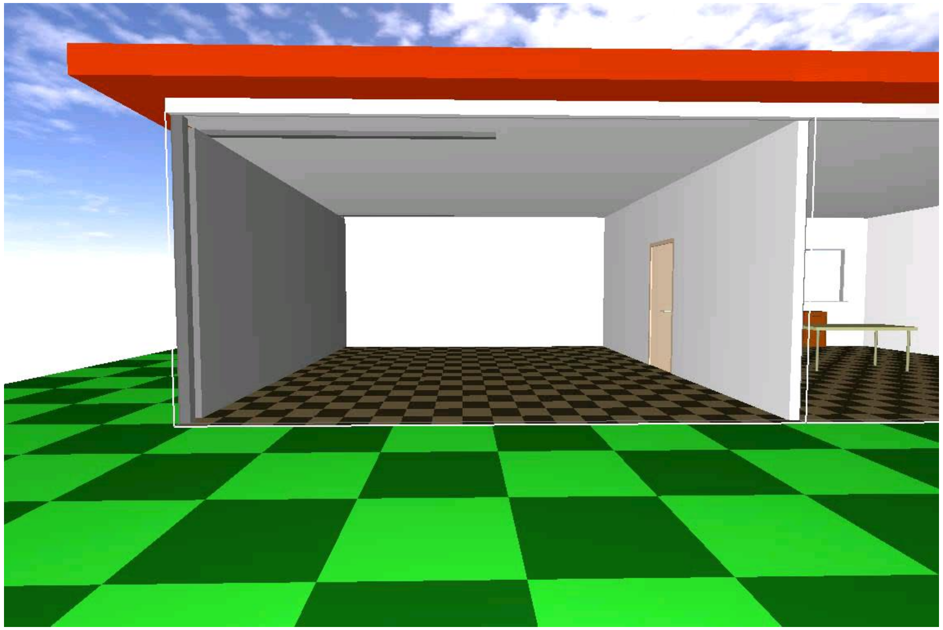
The Non-parametric Model



The Non-parametric Model

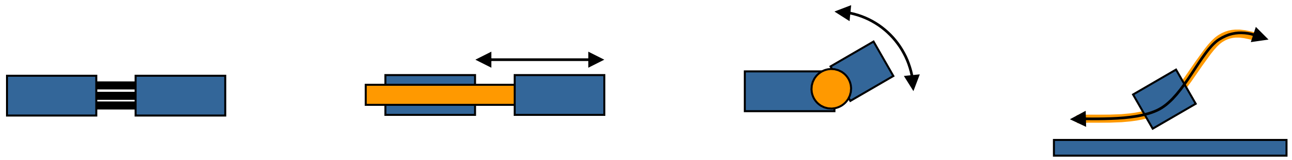
- Articulated objects have few DOF
- Articulated parts move on low-dimensional manifold
- Recover manifold + learn transformation





Which model is the best?

- Four candidate models



- More general models always fit
- Simpler models are more robust

Step 2: Model Selection

- Bayesian theory: Compare model posteriors

$$p(\mathcal{M} \mid \mathcal{D}) = \int \frac{p(\mathcal{D} \mid \mathcal{M}, \theta) p(\theta \mid \mathcal{M}) p(\mathcal{M})}{p(\mathcal{D})} \mathrm{d}\theta$$

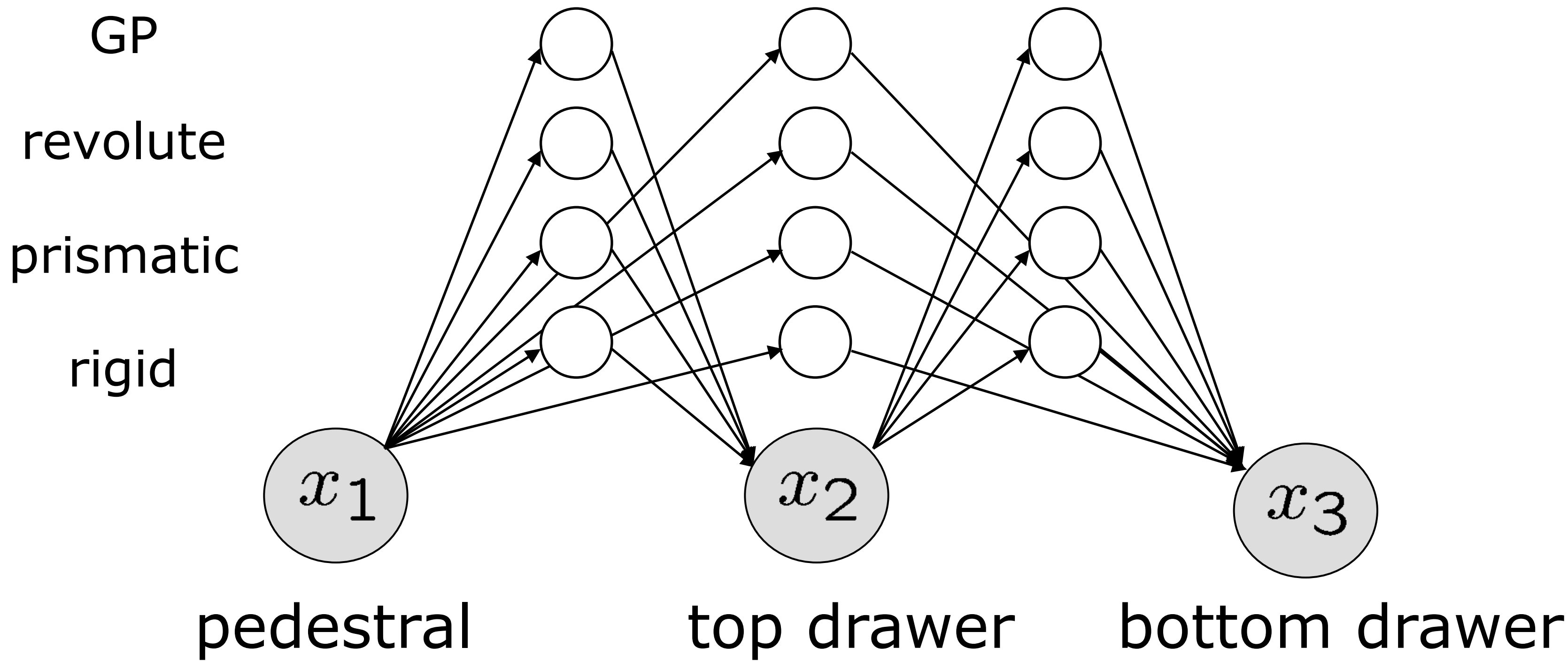
- This integral can be approximated using the Bayesian Information Criterion (BIC)

$$\mathrm{BIC} = \underbrace{-2 \log p(\mathcal{D} \mid \mathcal{M}, \hat{\theta})}_{\text{data likelihood}} + \underbrace{k \log n}_{\text{model complexity penalty}}$$

[Sturm et al., IJCAI'09]

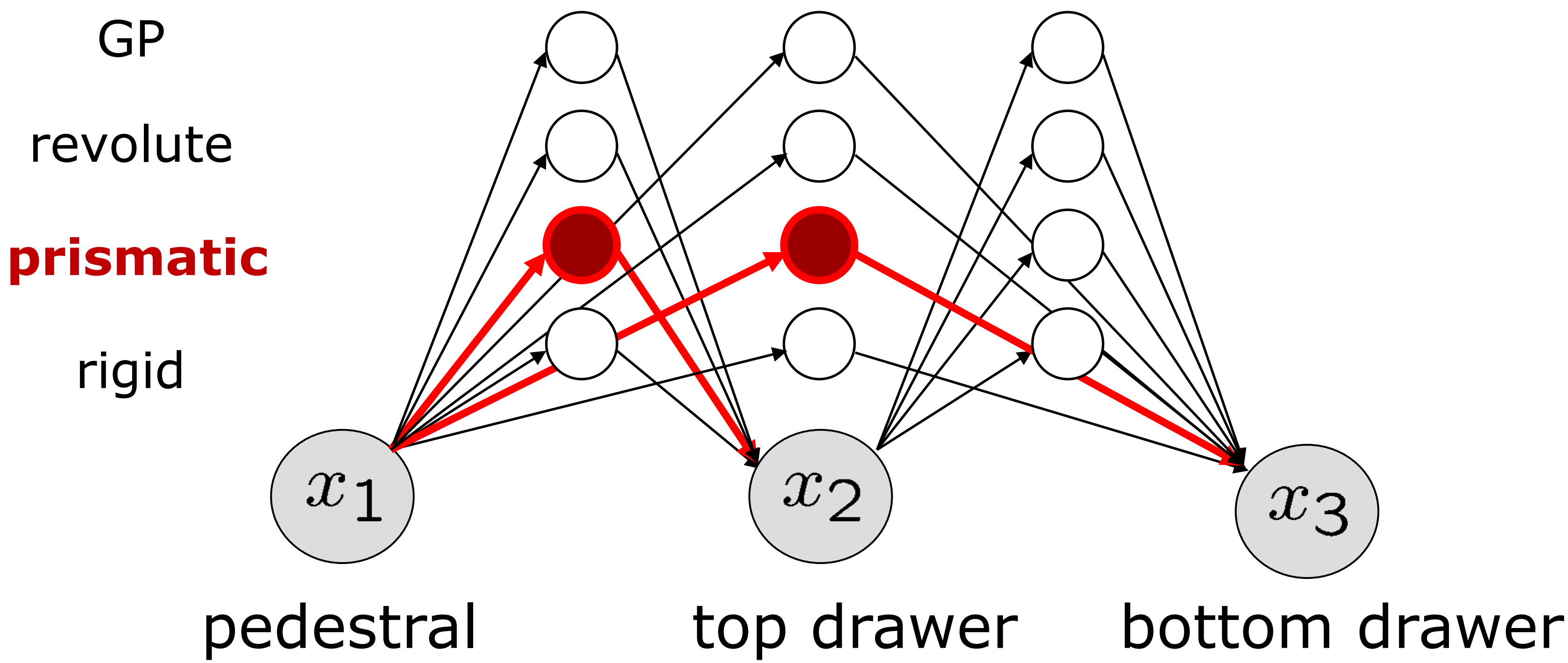
Inferring the Topology

- Find best kinematic tree (no loops)
- Model as a graph, use BIC as edge cost
- Minimum spanning tree is optimal solution



Inferring the Topology

- Find best kinematic tree (no loops)
- Model as a graph, use BIC as edge cost
- Minimum spanning tree is optimal solution



[Sturm et al., IJCAI'09]

Experiment: Microwave Oven

Input sequence

[Sturm et al., IJCAI'09]

Microwave Oven: Learned Model

Graphical Model	Kinematic Function	Reprojection of Learned Model
		

[Sturm et al., IJCAI'09]

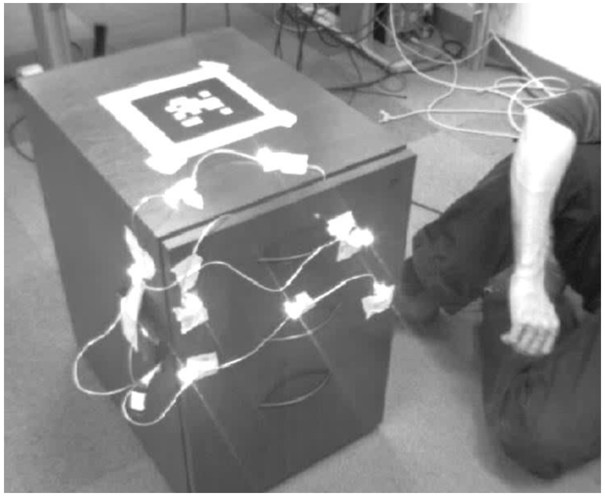
Office Pedestral: Learned Model

Learned Graphical Model

Kinematic Function of Top Drawer

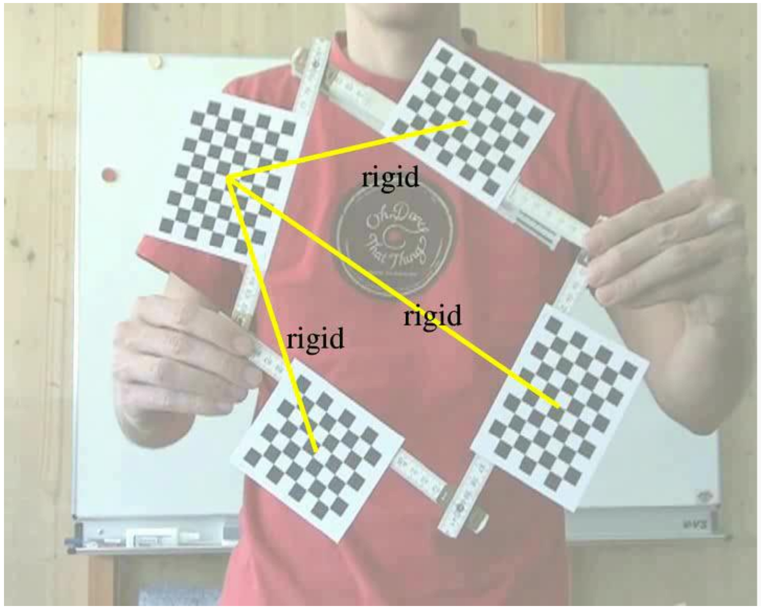
Kinematic Function of Bottom Drawer

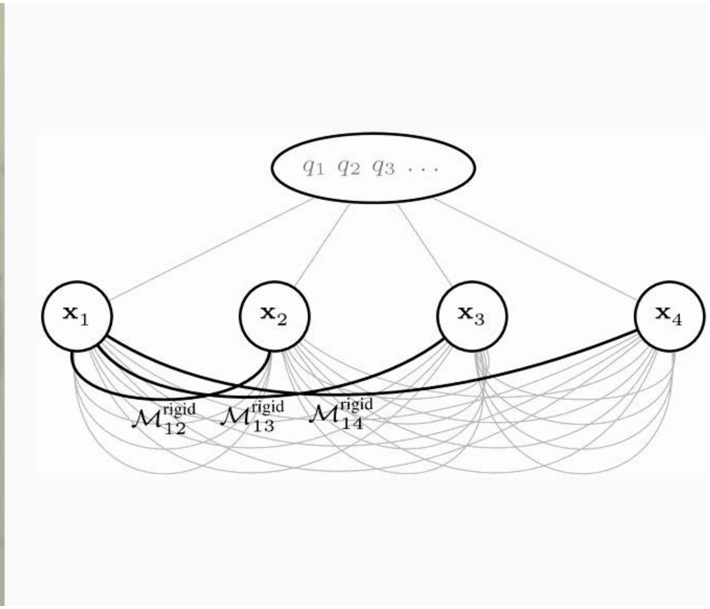
Reprojection of Learned Model



Closed Kinematic Chain

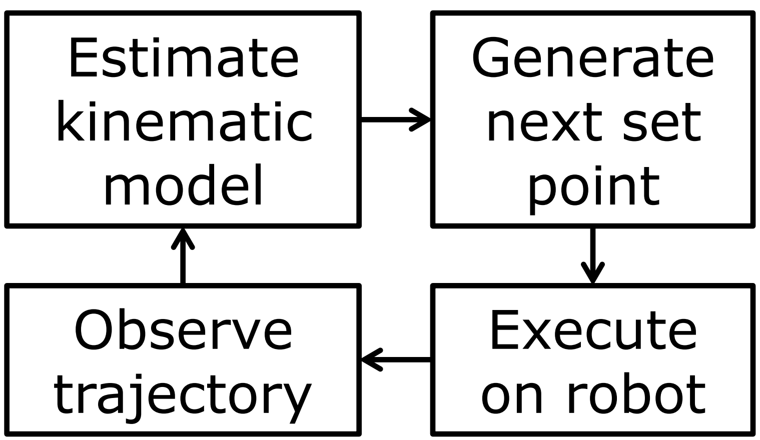
- Approach can be generalized to arbitrary kinematic graphs (including loops)
- Estimate the DoF of the system
- Significantly increased complexity





Operating Articulated Objects

- **Closed-loop** model estimation and control (joint encoders)
- Learn kinematic model during execution
- Improved accuracy through repeated interactions

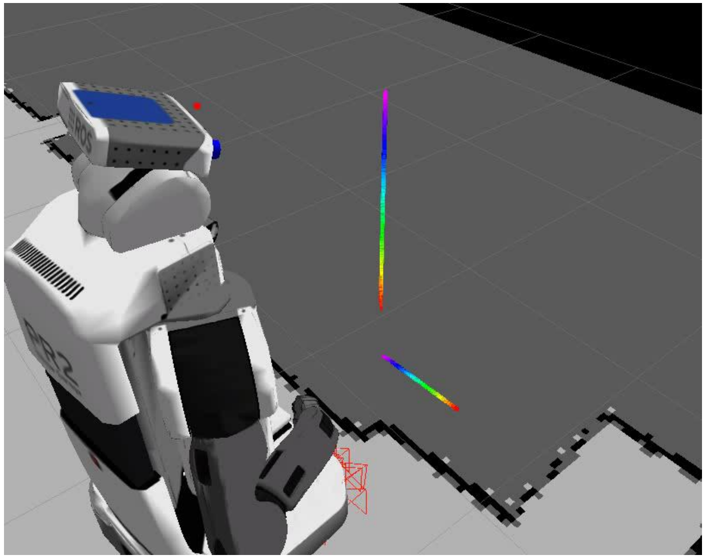


Georgia Tech

[ICRA'12]

Towards Autonomous Mapping of Articulated Objects

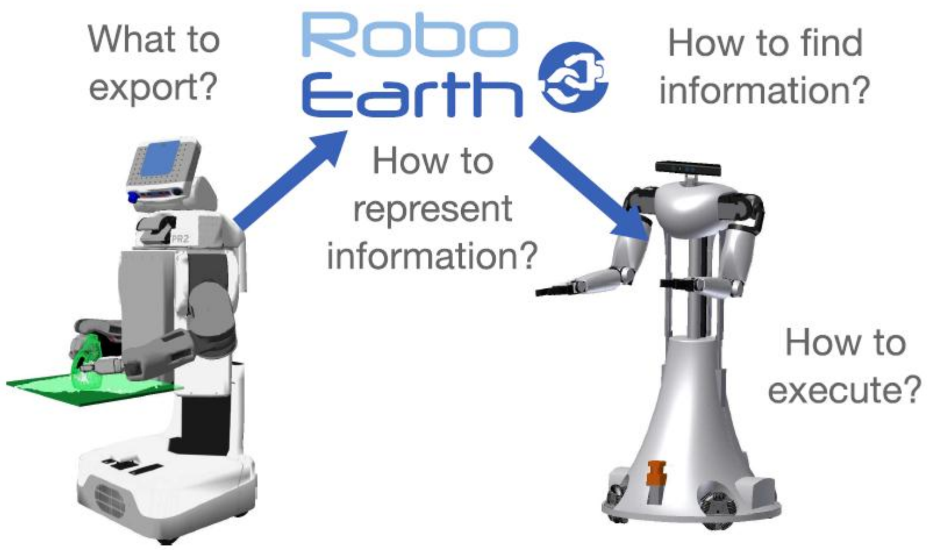
- **Visual perception** + closed-loop model estimation and control
- Store/retrieve models in the map



Technical University of Munich

Towards Autonomous Mapping of Articulated Objects

RoboEarth project (FP7): store/retrieve models in a world-wide data base, exchange with other robots



Eindhoven University of Technology, Philips Innovation Services, University of Stuttgart, Swiss Federal Institute of Technology Zurich, University of Zaragoza, Technische Universität München

Conclusions

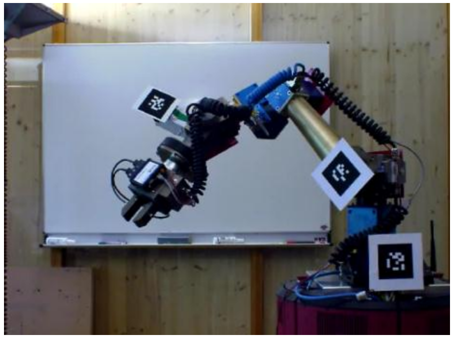
- Integrated Bayesian framework for modeling articulated objects
- Fully available as open-source
- Significantly increases the flexibility of service robots in unstructured environments
- Actively used by several independent research groups and research projects

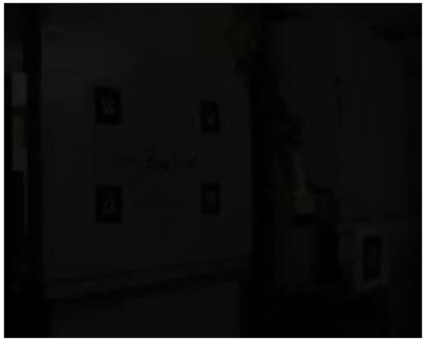
PhD Thesis: “Approaches to Probabilistic Model Learning”

- Chapter 3: Body schema learning
  [ICRA’08, RSS’08, JP’09, GWR’09]
- **Chapter 4+5: Articulated objects** ⇐ **this talk**
  [IJCAI’09, ICRA’10,IROS’10, RSS’10,JAIR’11]
- Chapter 6+7: Tactile sensing
  [IROS’09, IROS’10, TRO’11]
- Chapter 8: Imitation learning
  [ICRA’09]

3 journal articles, 14 conference and workshop papers, h-index 8, >160 citations

Thank You For Your Attention!





Many thanks go to:

Wolfram Burgard, Kurt Konolige, Cyrill Stachniss, Christian Plagemann and all members of the AIS lab in Freiburg!

Future Work

Research Projects

- First-MM (EU FP7)
  Learn flexible manipulation skills
- RoboEarth (EU FP7)
  Exchange models between robots
- A8 Project in SFB/TR8 (DFG)
  Apply to humanoid robots
- TidyUp Robot Project (Willow Garage)
  Generalized mapping

Research Groups

- U Freiburg, Autonomous Intelligent Systems **[Cyrill Stachniss, Wolfram Burgard]**, Humanoids Lab **[Maren Bennewitz]**
- TU Eindhoven, Mechanical Engineering
  [Rob Janssen, Marinus van de Molengraft]
- TU Munich, Autonomous Intelligent Systems
  [Thomas Rühr, Dejan Pangercic, Michael Beetz]
- ETH Zurich, Dynamic Systems and Control
  [Ramos de la Flor, Nico Hübel, Rafaello D'Andrea]
- FZI Karlsruhe, Intelligent Systems and Product Engineering
  [Andreas Hermann, Rüdiger Dillmann]
- Bonn-Rhine-Sieg University, b-it-bots
  [Jan Paulus, Nico Hochgeschwender, Gerhard Kraetzschmar]
- Georgia Tech, Healthcare Robotics Lab
  [Advait Jain, Charlie Kemp]

Future Work: Flying Manipulation

- Quadcopters
  - 100g: smartphone or video camera(s)
  - 500g: Kinect, gripper, dual core processor
  - 2kg: more advanced sensors, whole laptop, actuated manipulator, carry heavier objects
- Applications
  - 3D mapping and navigation
  - Flying consumer cameras (ski, hiking,…)
  - Tidy up tasks (return empty beer bottles to crate)

Future Work: 3D Perception

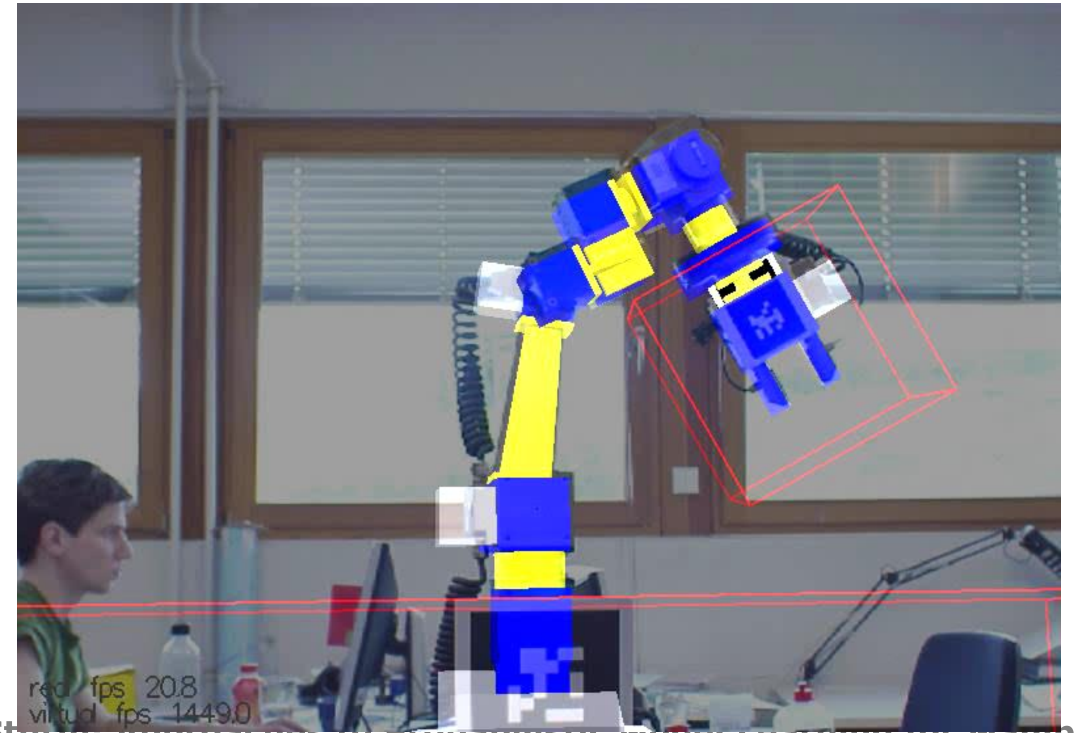
- 3D tracking, localization and mapping
  - Dense methods
  - Convex optimization
  - 3D reconstruction
- Active perception (using robots)
  - Active segmentation
  - Visual navigation with quadcopters
  - Flying manipulation
- Benchmarking

Body Schema Learning

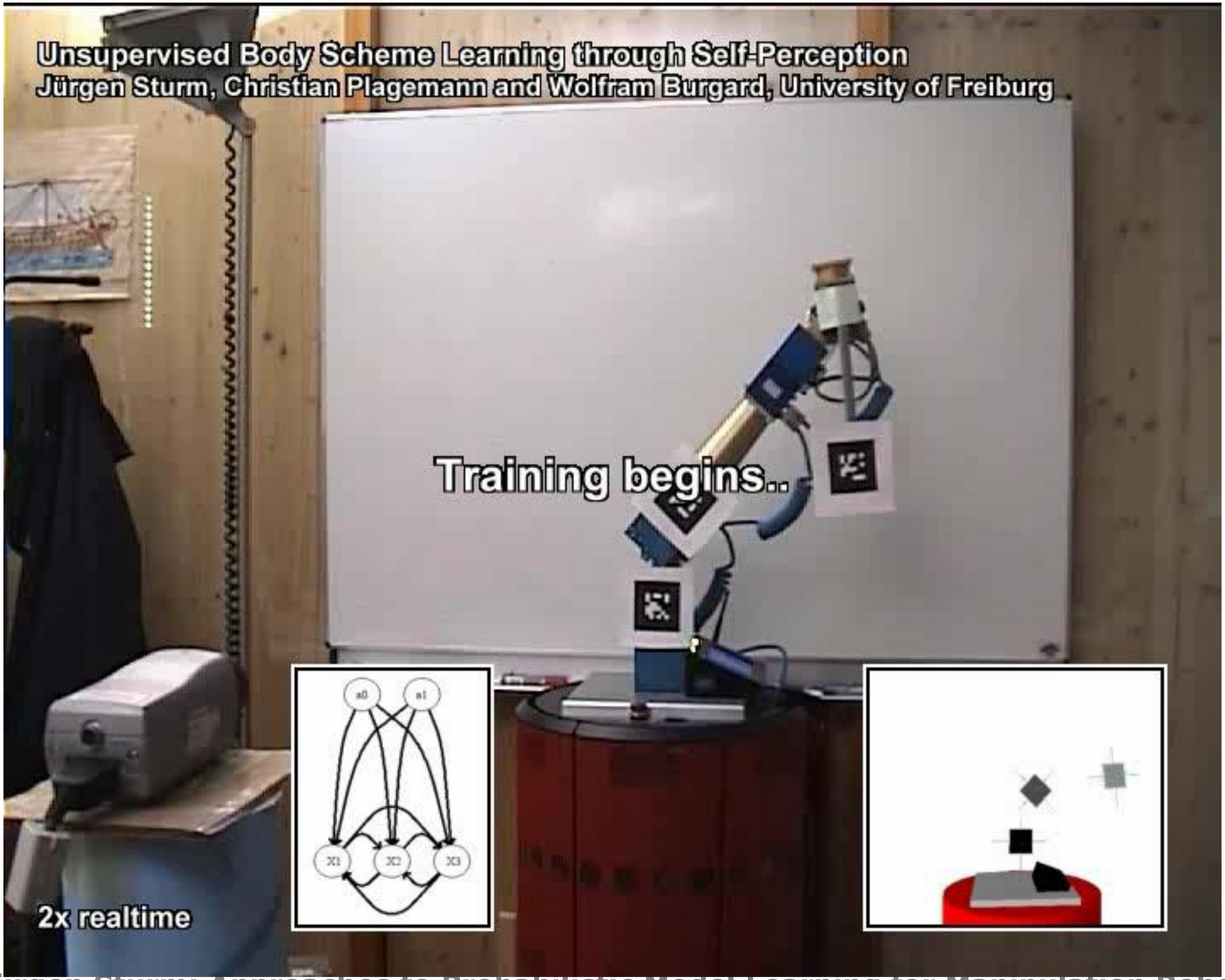
Motivation

Existing robot models are typically

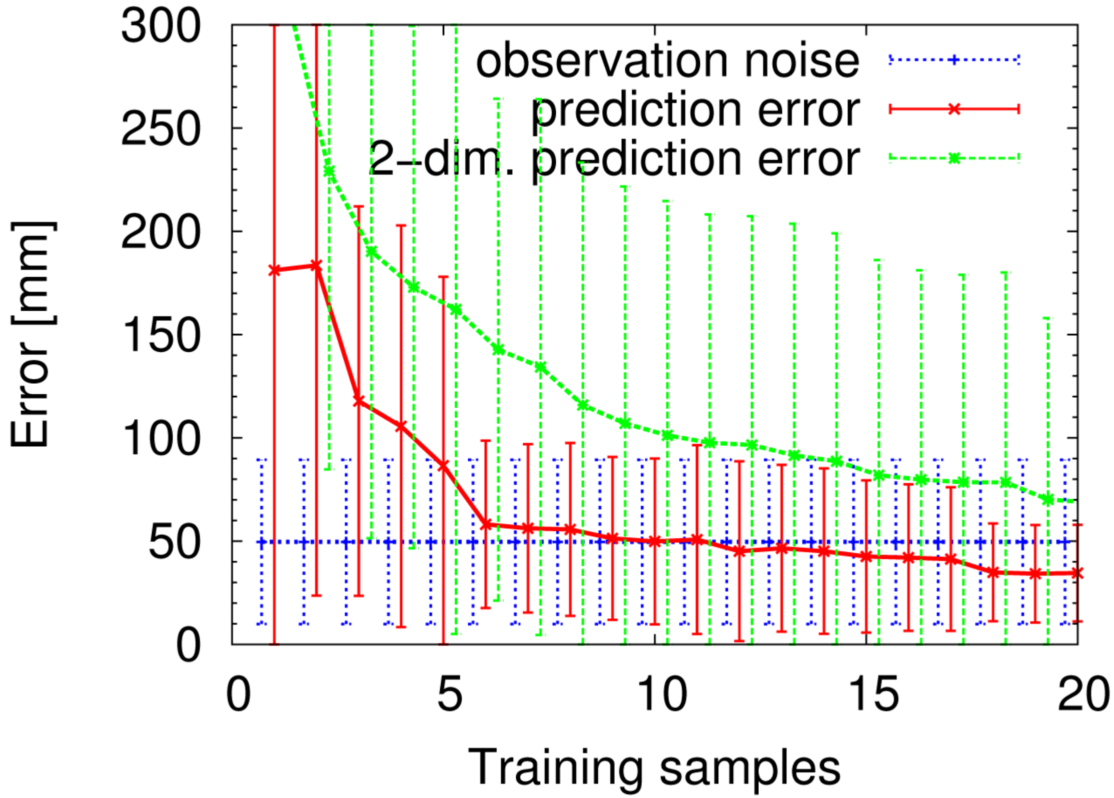
- specified (geometrically) in advance and the
- parameters are calibrated manually



Experiments

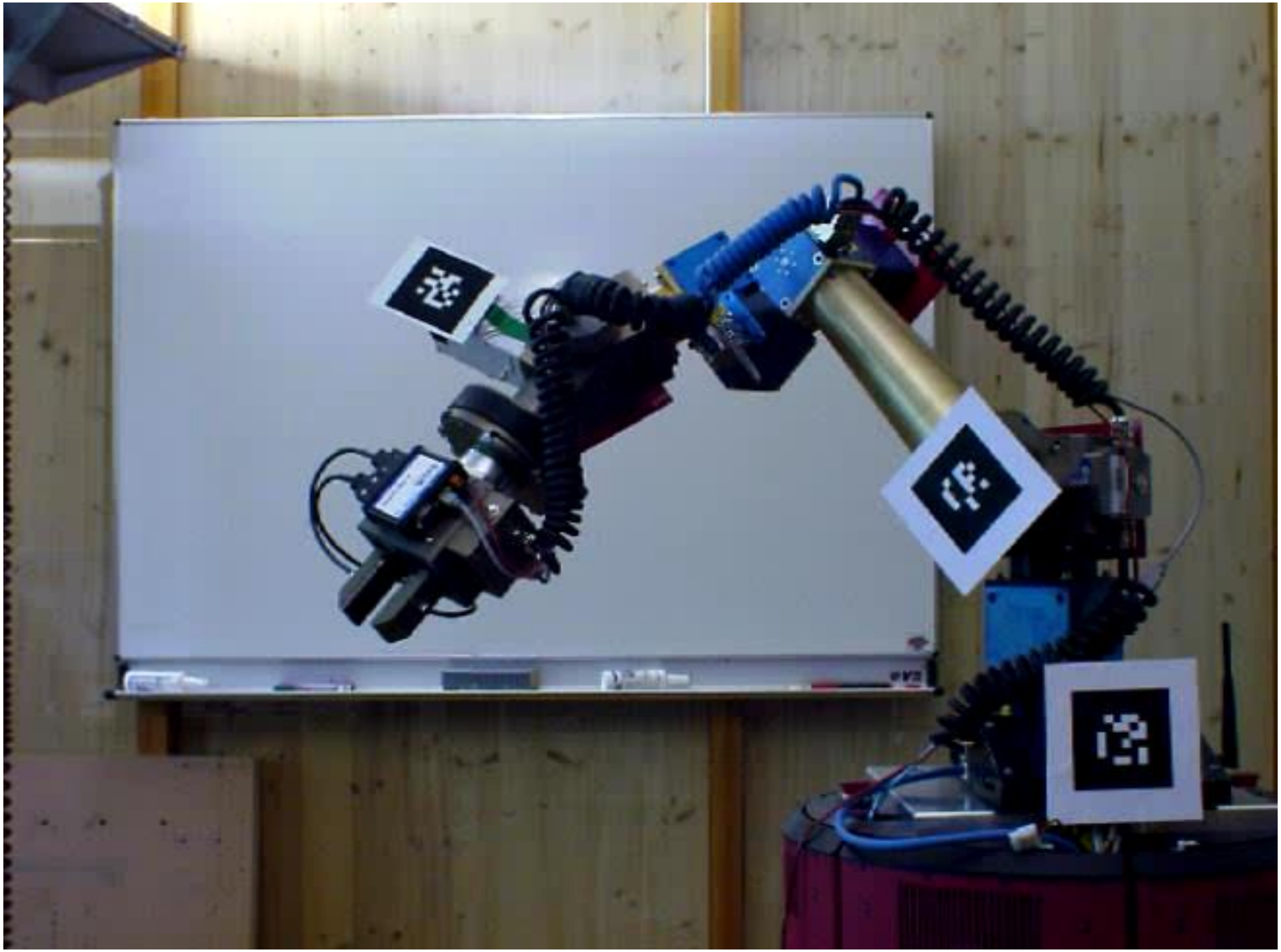


Evaluation: Forward Kinematics



- Fast convergence (approx. 10-20 iterations)
- High accuracy (higher than direct perception)

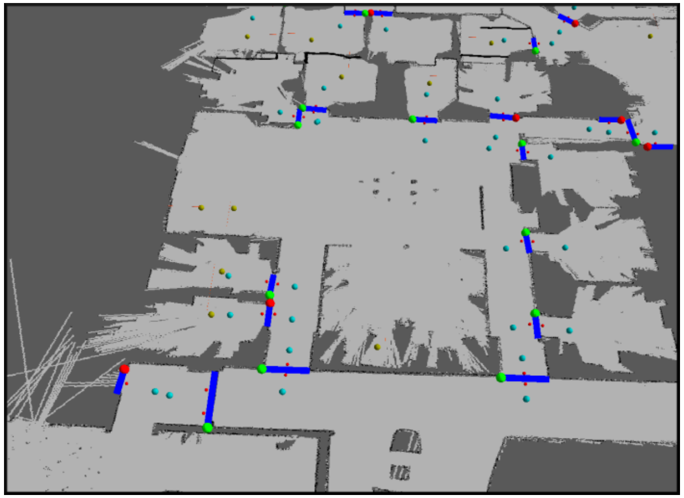
Life-long Adaptation



Articulated Objects

Related Work (1)

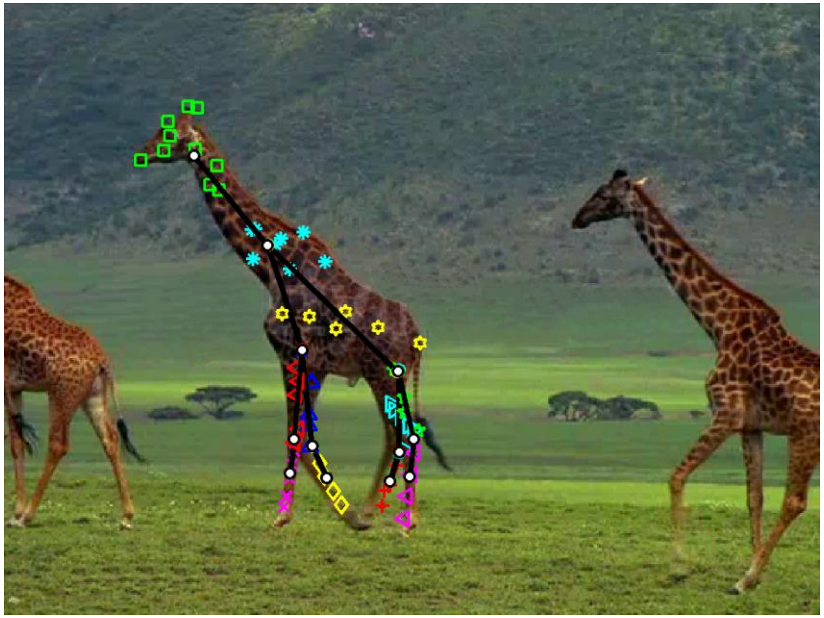
- Door and door handle detection
- Robust control
- Door locations specified in map
- Scripted turn and push motion



[Meeussen, Wise, Glaser, Chitta, McGann, Mihelich, Marder-Eppstein, Muja, Eruhimov, Foote, Hsu, Rusu, Marthi, Bradski, Konolige, Gerkey, Berger, ICRA 2009]

Related Work (2)

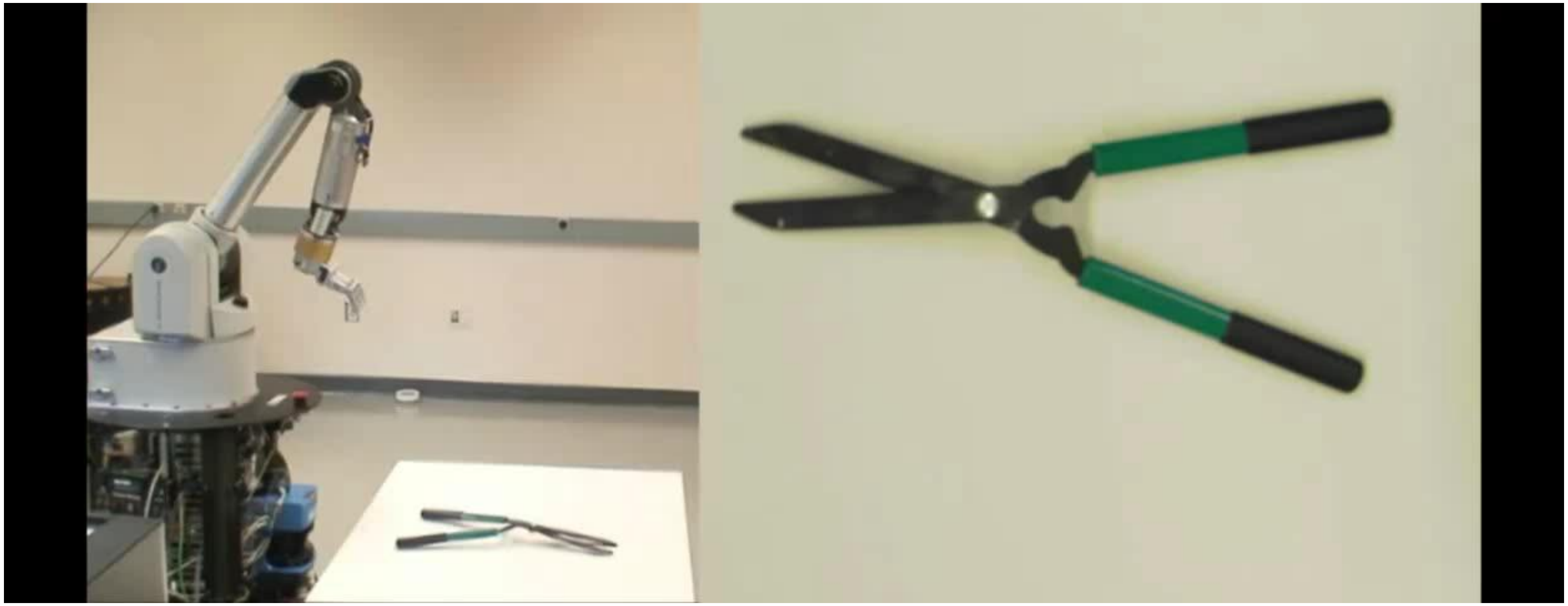
- Motion Capture and Video
- 2D/3D Feature Tracks
- Recover stick figures
- Learns graphical model



[Ross, Tarlow and Zemel, IJCV 2010]

Related Work (3)

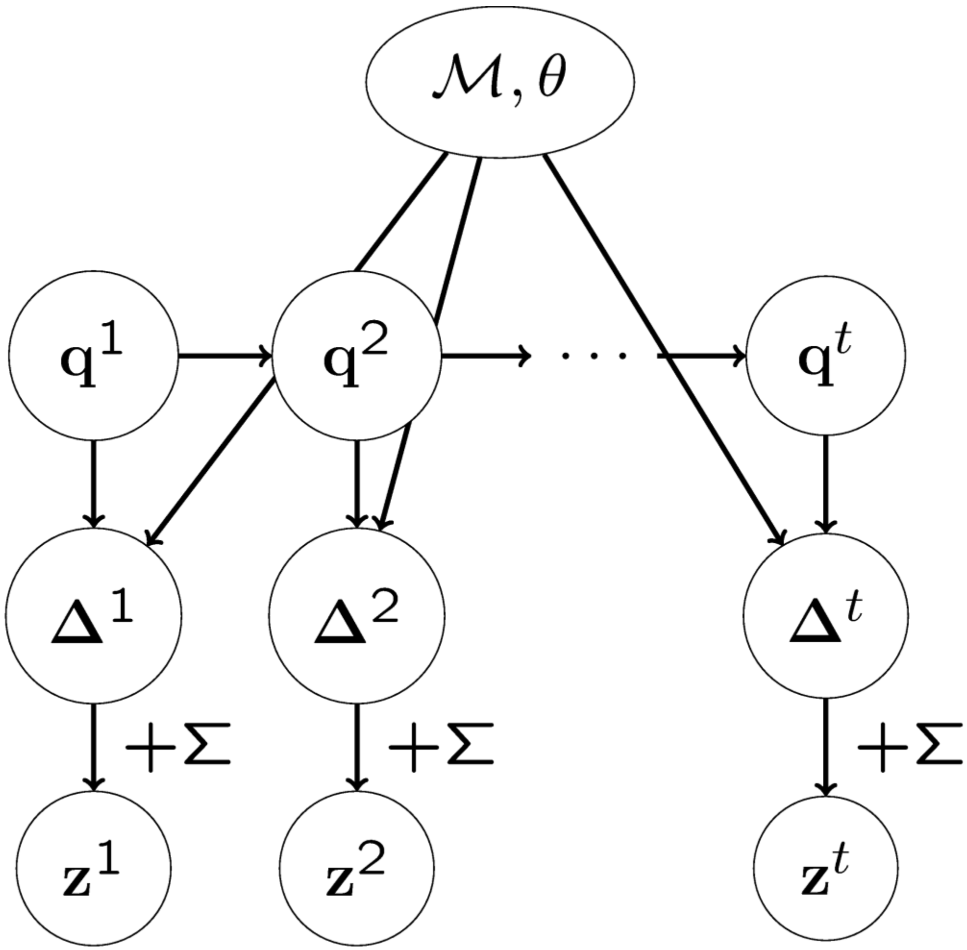
- Manipulator + Camera
- Interactive Perception
- Tracks KLT-Features
- Min-cut algorithm on feature graph



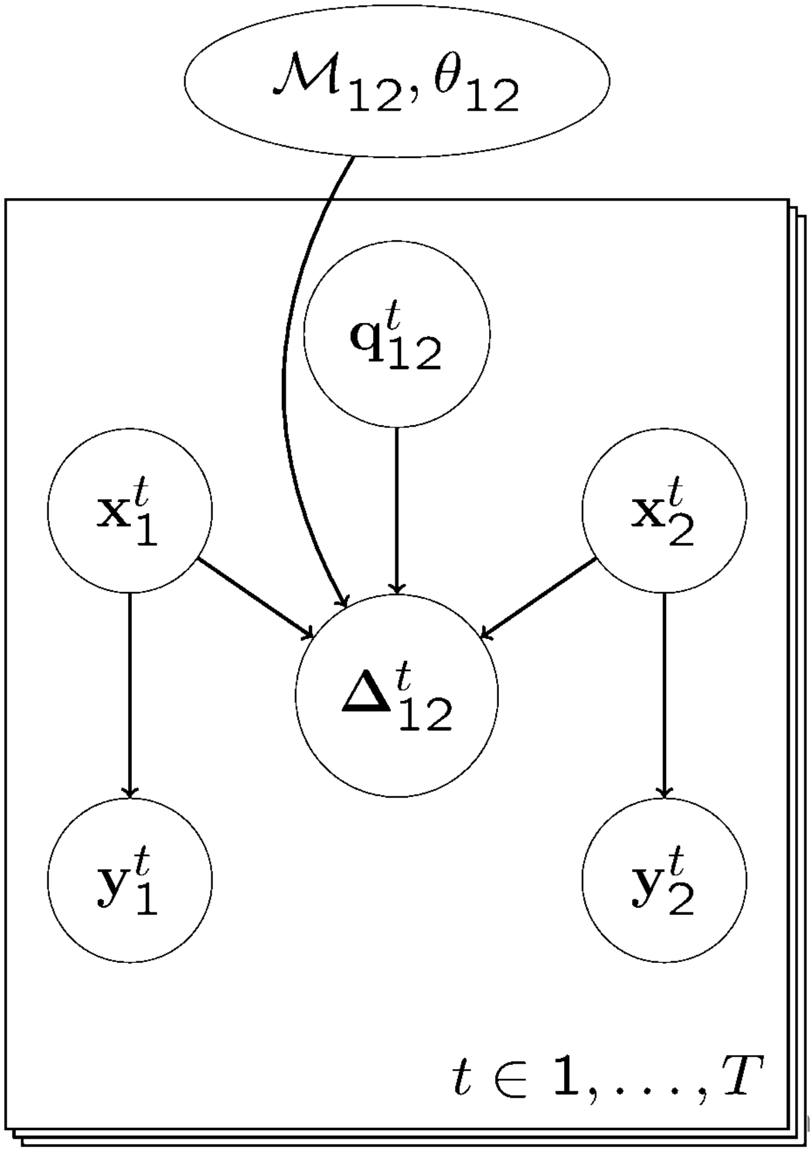
[Katz and Brock, RSS 2008]

Process Model

- Kinematic model
- Configuration
- True pose
- Observed pose

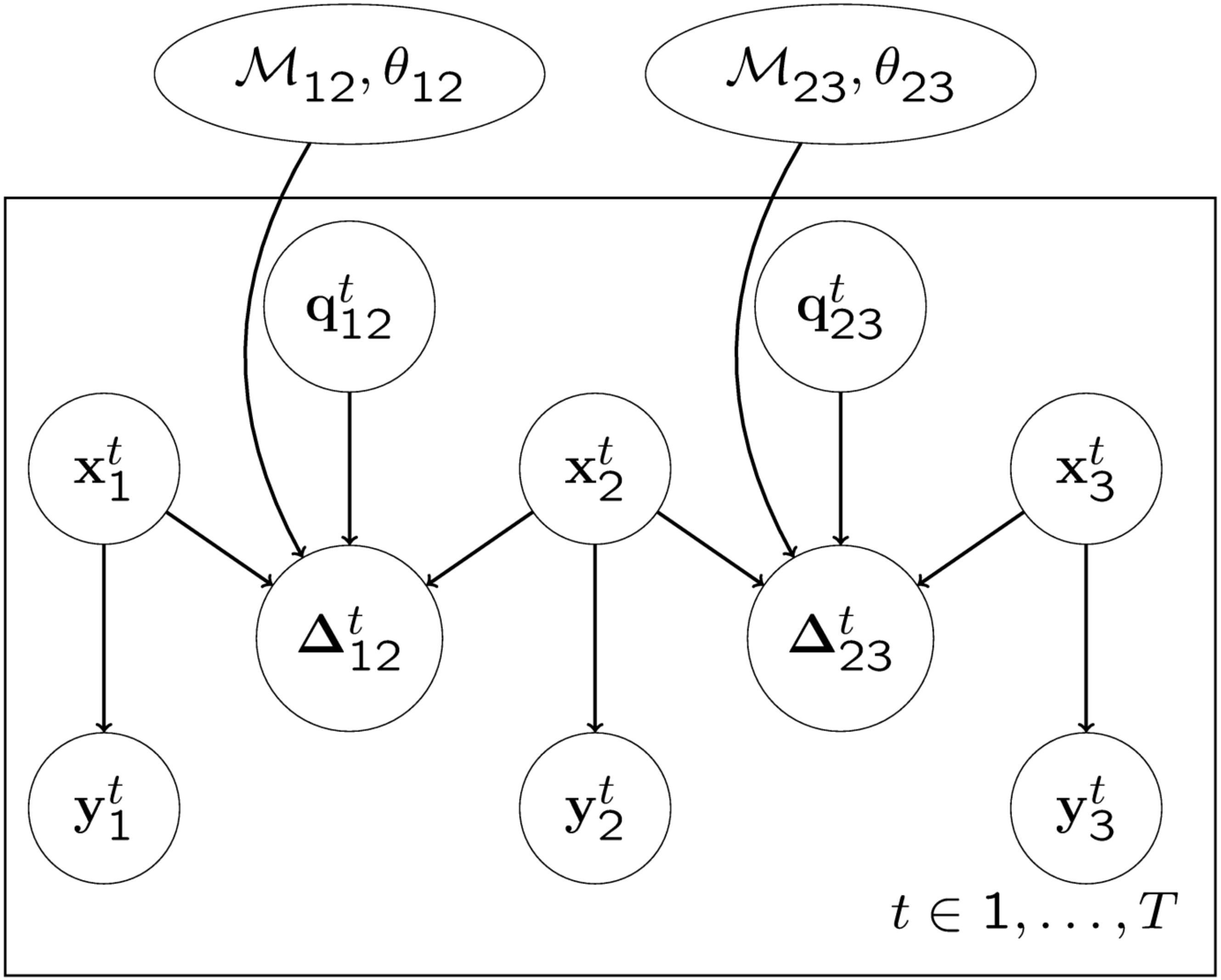


Process Model for 2 parts

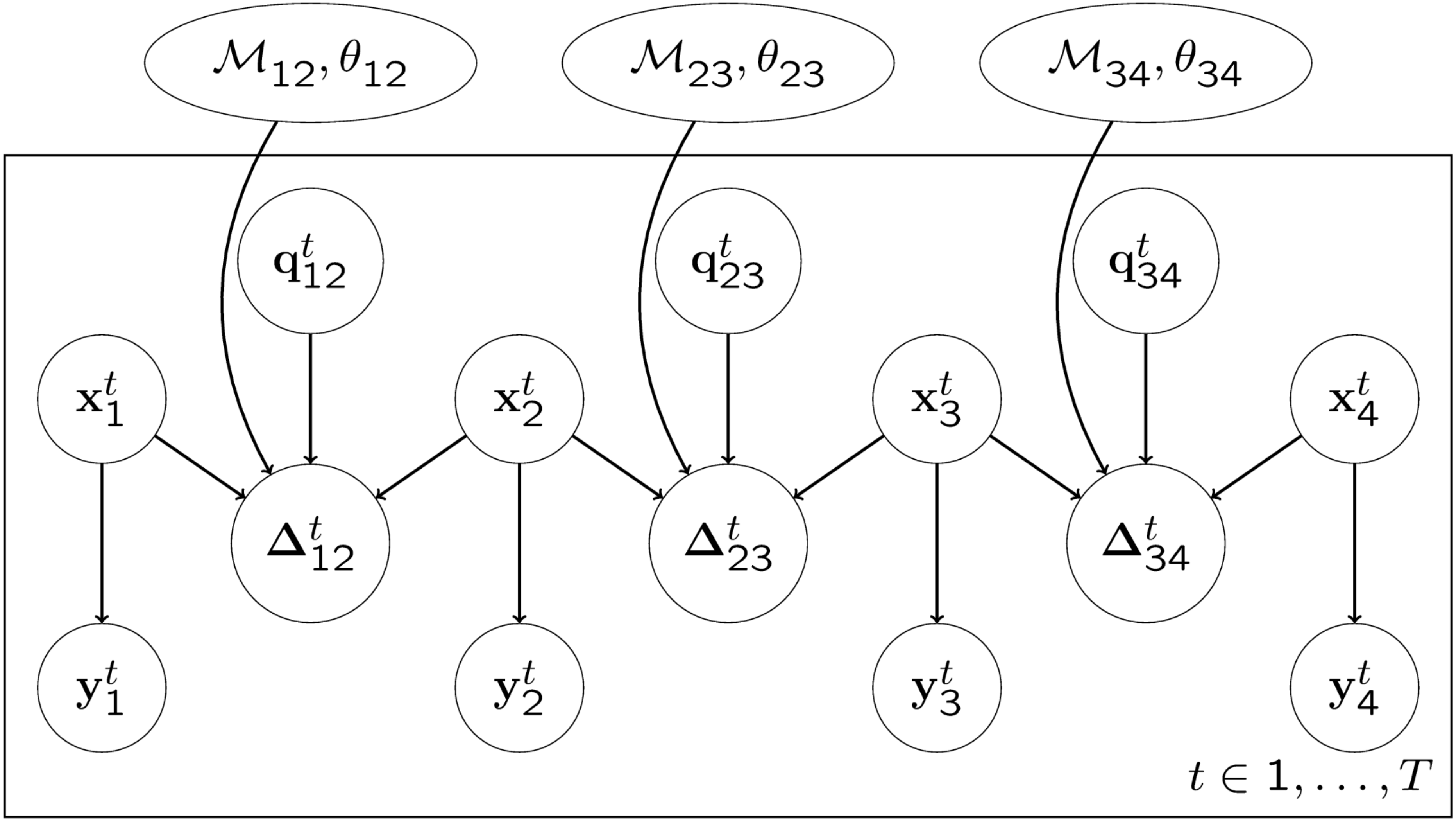


- Kinematic model
- Configuration
- True poses
- True transformation
- Observed poses

Process Model for 3-chain



Process Model for 4-chain



Examples 1/3

fridge

drawer

Examples 2/3

dishwasher

.. and tray

Examples 3/3

water tap

valve of a radiator

Model Clustering

- Given two observed trajectories, should we select one or two models?

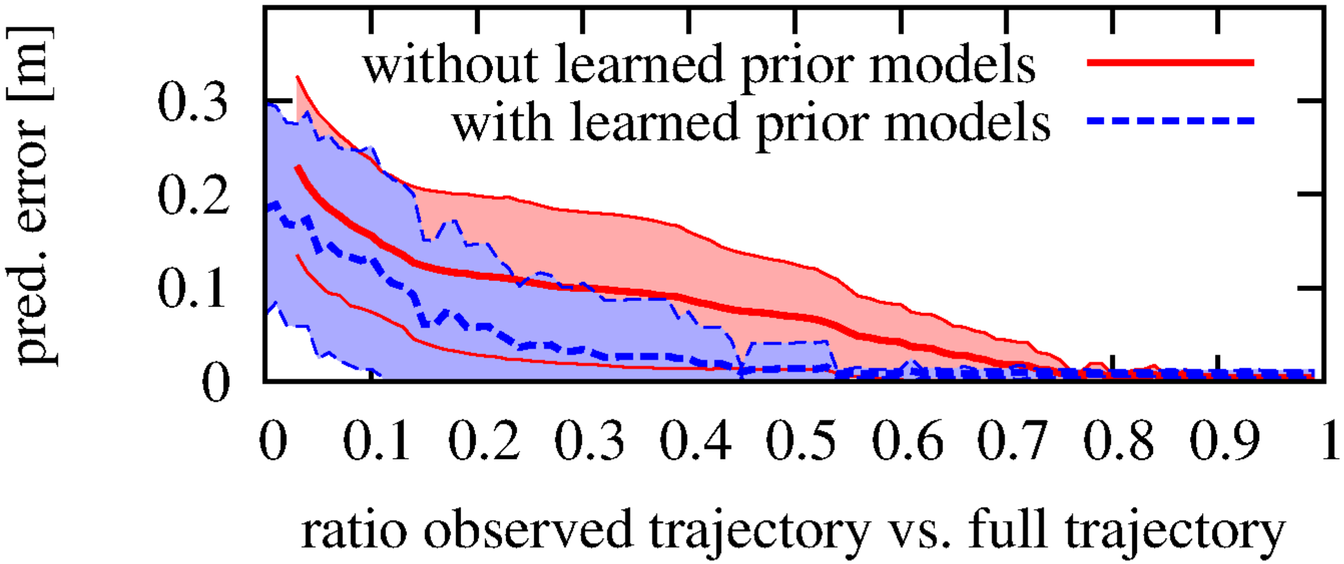
- Bayesian model comparison

If $$p(\mathcal{M}_{1+2} \mid \mathcal{D}) > p(\mathcal{M}_1, \mathcal{M}_2 \mid \mathcal{D})$$

Then: Learn single model (single set of parameters but might fit data worse)

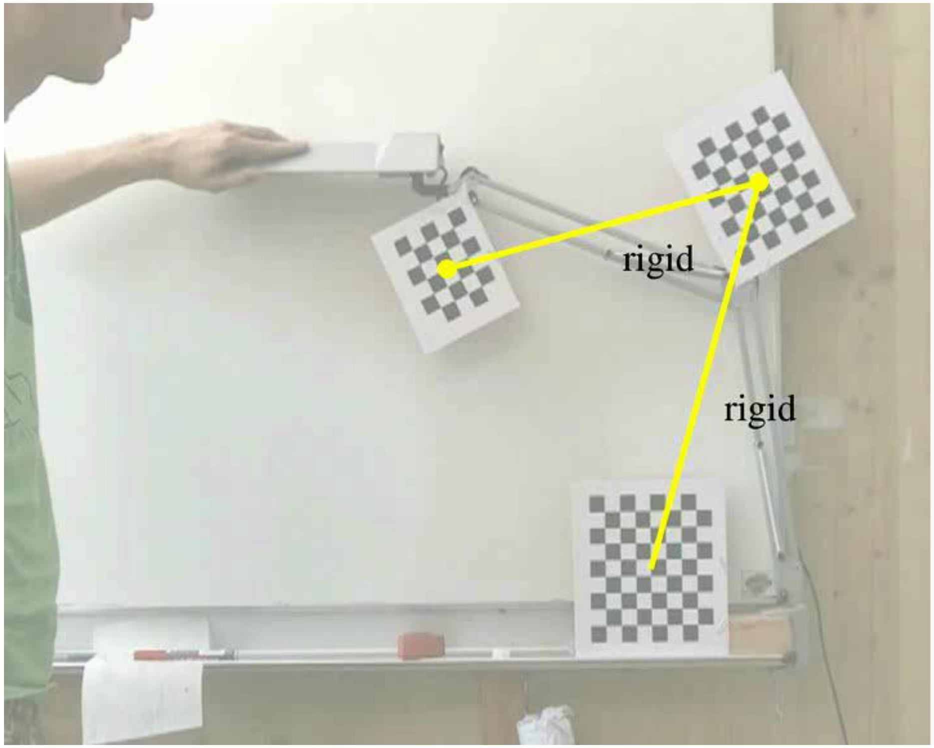
Else: Learn two models (double set of parameters but might fit data better)

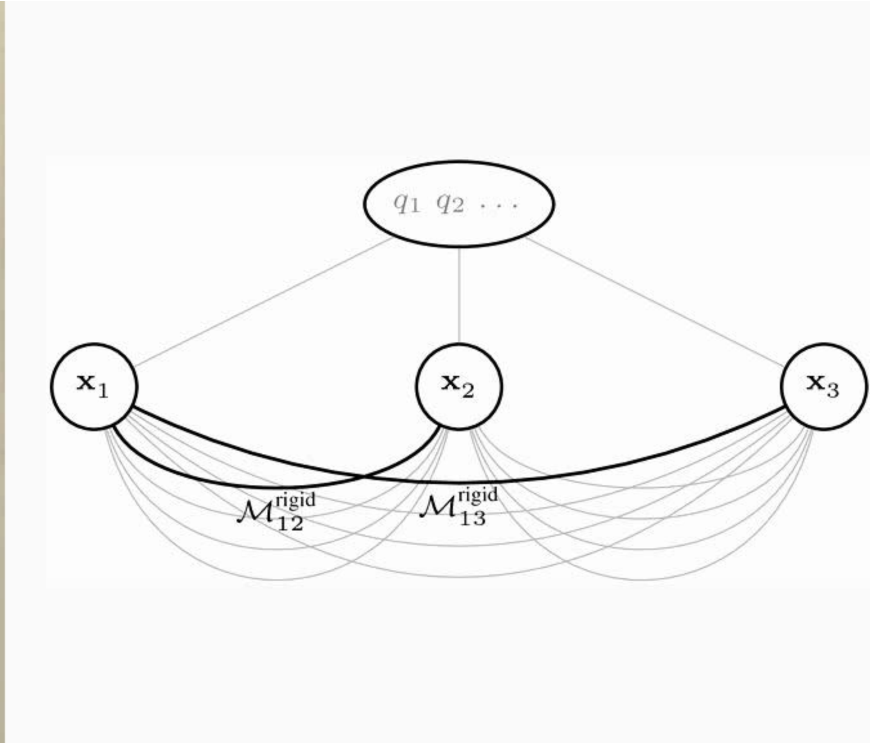
Exploiting Prior Information



- Using prior information significantly improves prediction accuracy

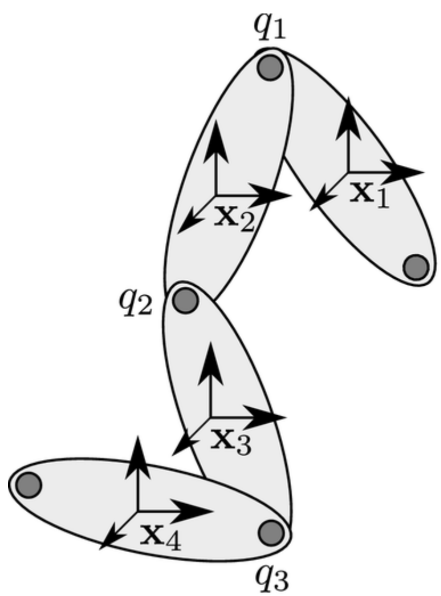
Example: Desk Lamp



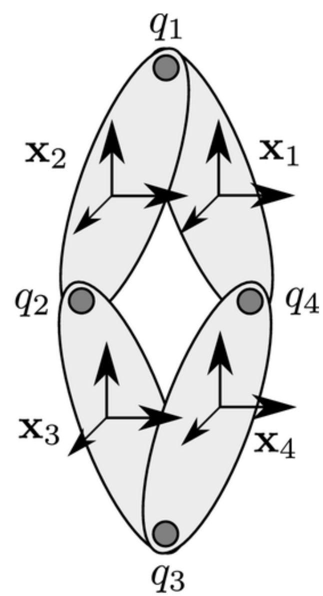


Estimate effective DOFs

- Closed chain objects might have less DOFs than the sum of their links

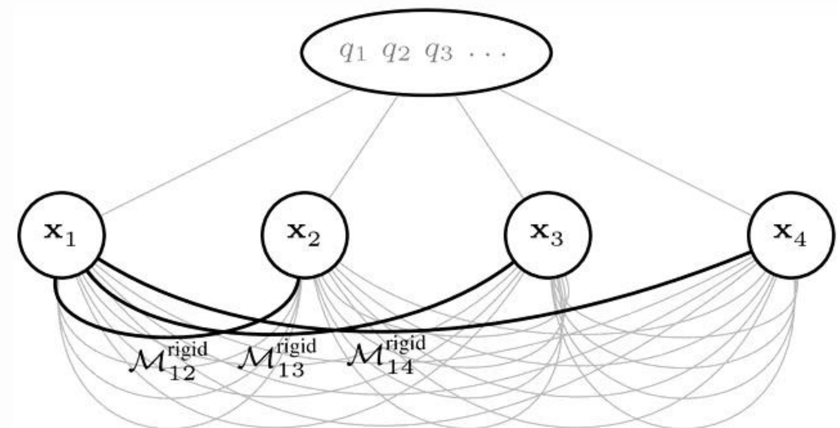


3 links
3 DOF

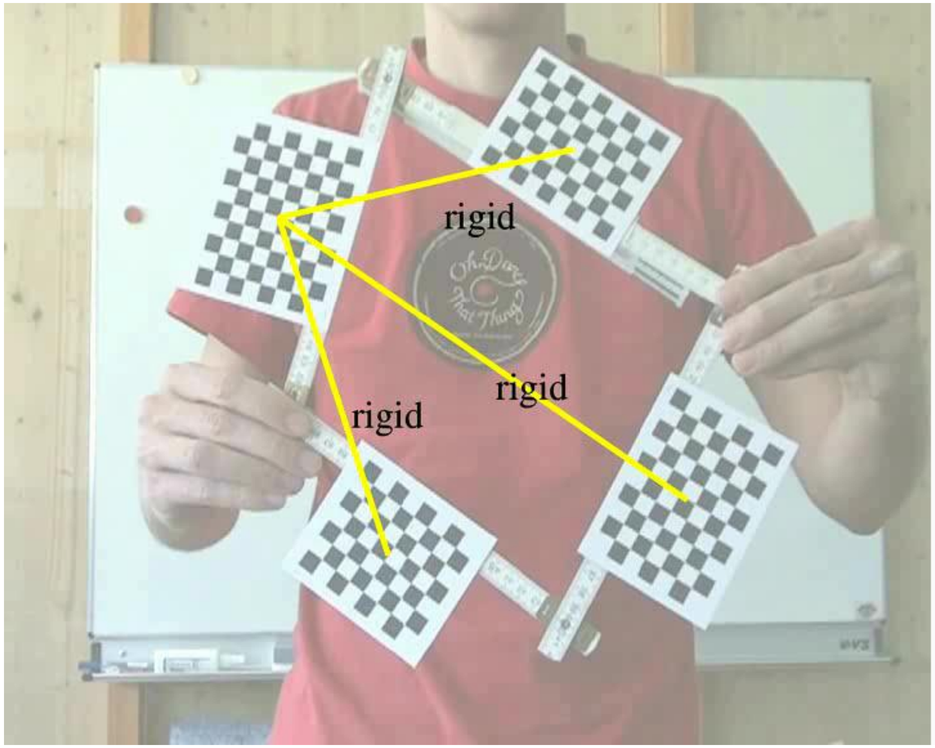


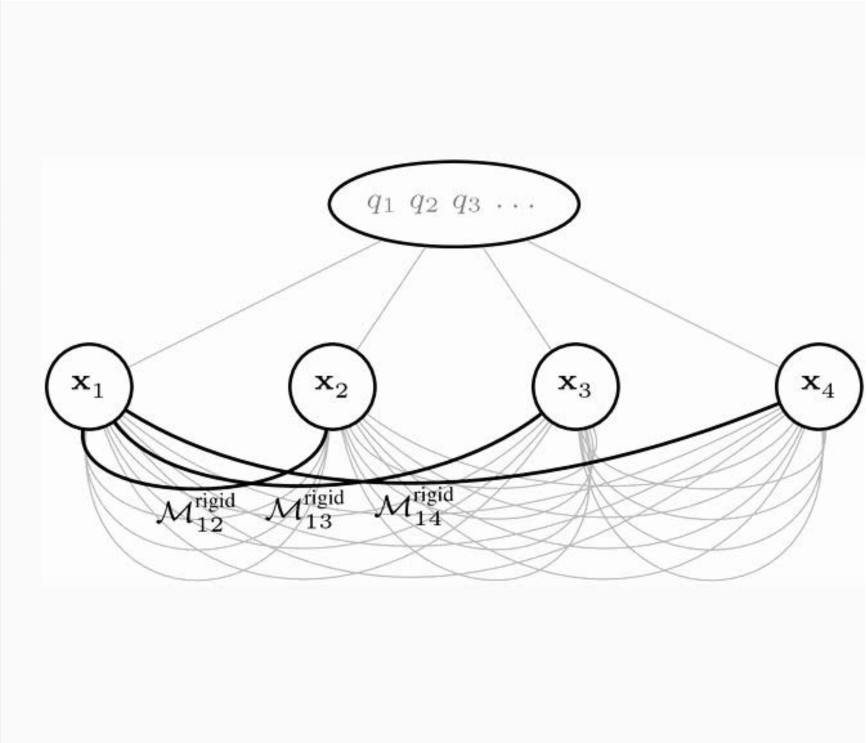
4 links
1 DOF

Example: Open Kinematic Chain

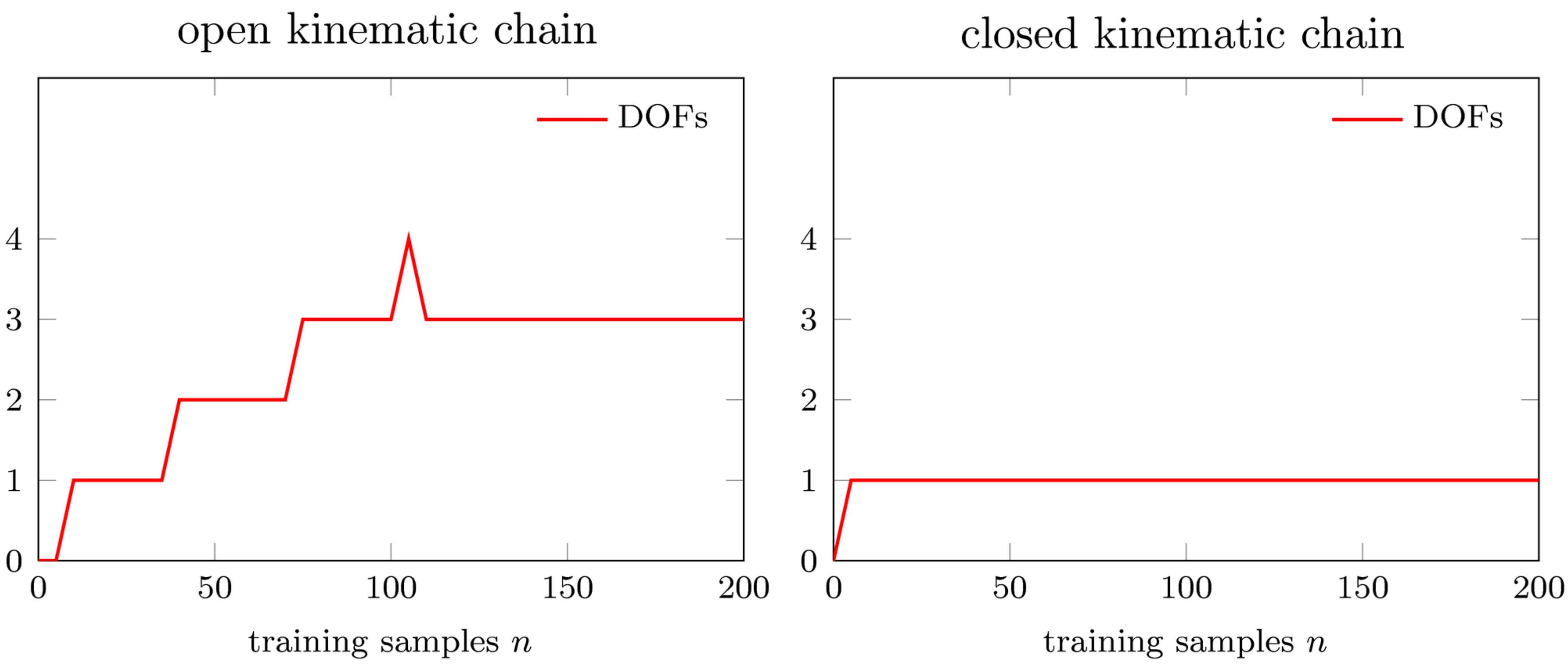


Example: Closed Kinematic Chain





Evaluation of DOFs



Marker-less Perception

- Artificial markers are not suitable for real-world applications…

➡ Can we learn the articulation models without using artificial markers?

 [ICRA'10]

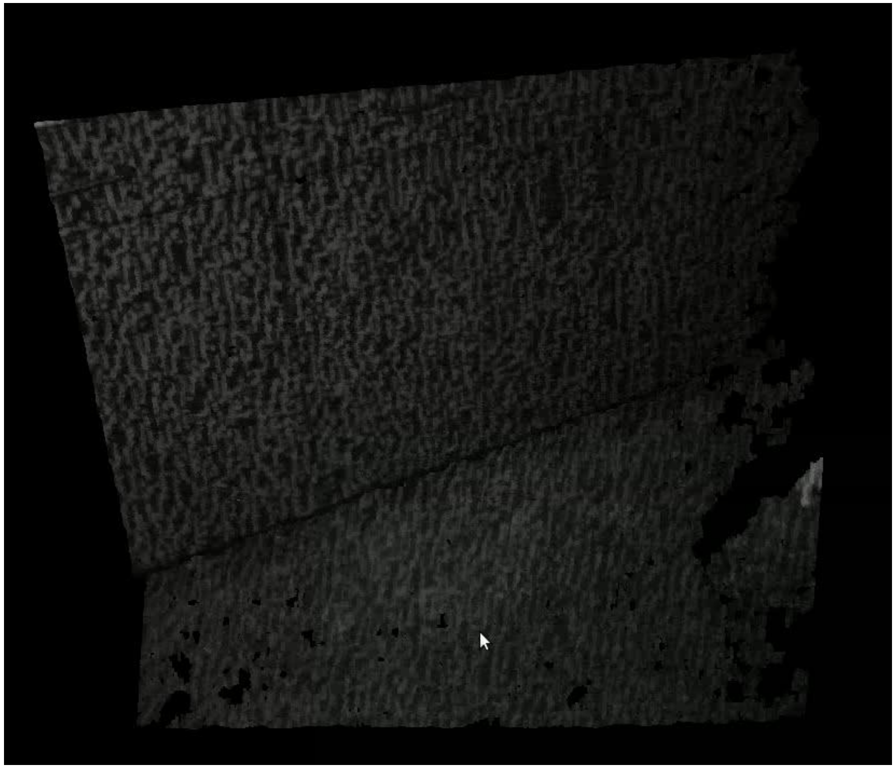
Marker-less Perception

- Detection and tracking of articulated objects in dense depth video
- Our approach: Plane segmentation and iterative pose fitting

 [ICRA'10]

Marker-less Perception

- Track detected objects
- Learn articulation models from observed trajectories



 [ICRA'10]

Marker-less Perception

- Track detected objects
- Learn articulation models from observed trajectories

Tactile Sensing

Example Data

- Robot grasps cup

left finger right finger

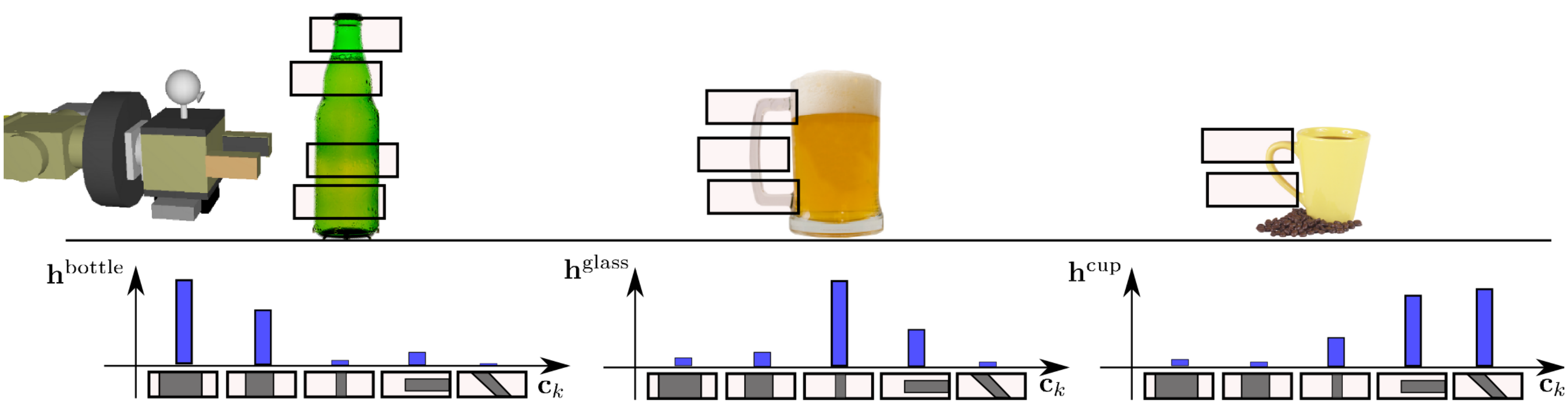
- Robot grasps pen

left finger right finger

Bag-of-Features Approach

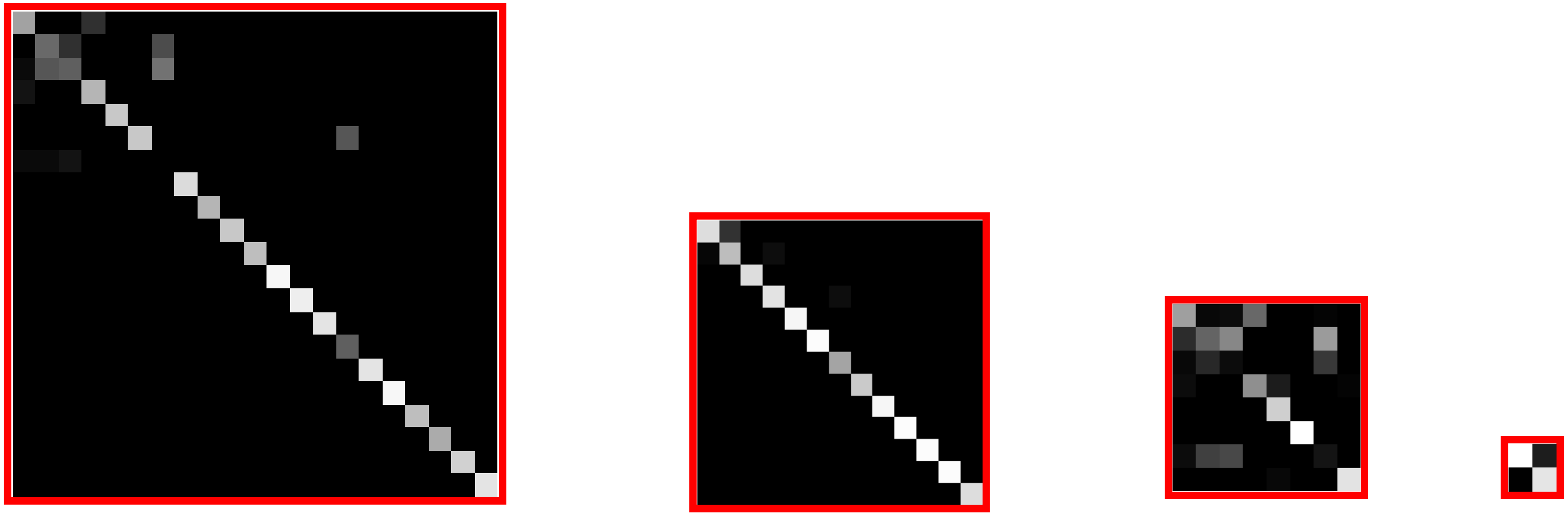
- Learn a codebook, i.e., a histogram relating features with object classes:

$$h_i^o \leftarrow h_i^o + \exp(-dist(\mathbf{c}_i, \mathbf{z})/l)$$



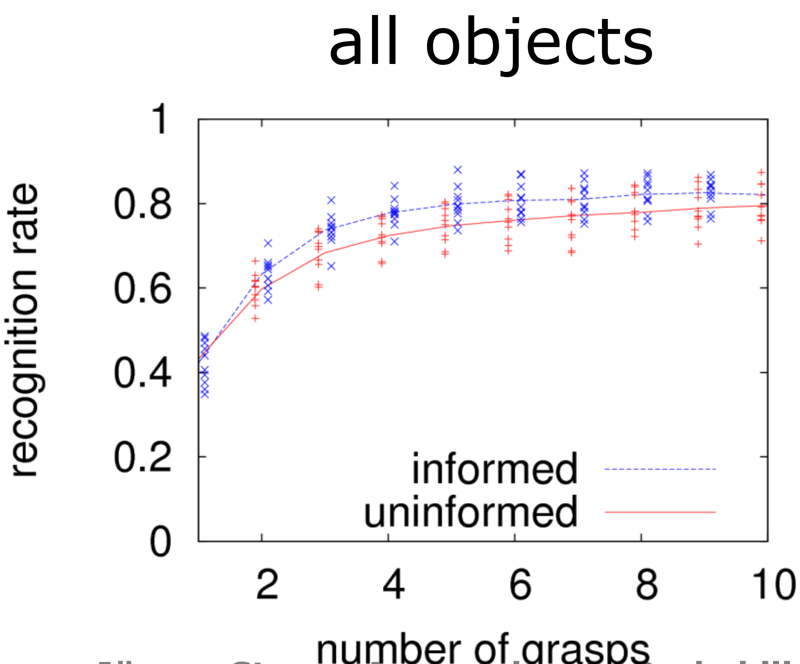
Recognition Rates

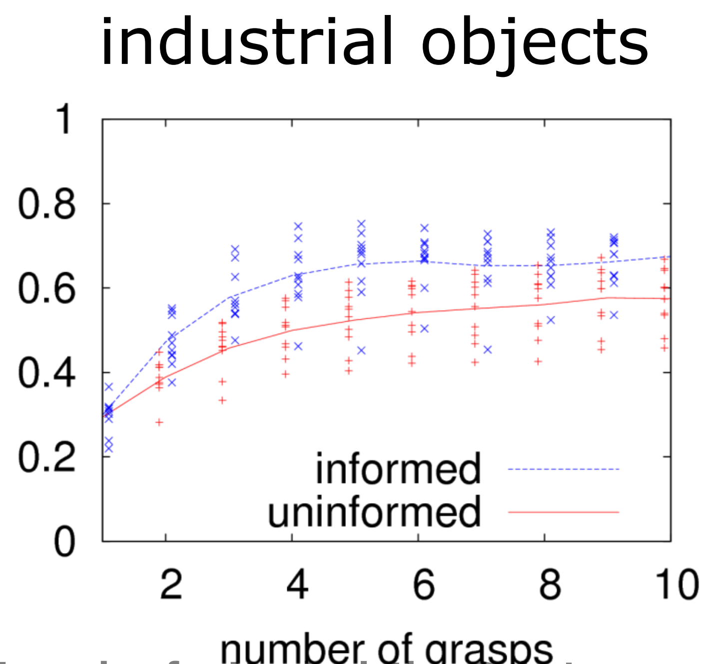
n	data set	recognition rate
21	all objects	84.4%
13	household objects	96.2%
8	industrial objects	58.0%
2	tennis balls	93.8%



Gain of Active Perception

- Significantly higher recognition rate (validated via t-test)
- More expressed for industrial objects (more difficult)





Motivation

- Visual information does not give the full picture about objects
  - *e.g.*

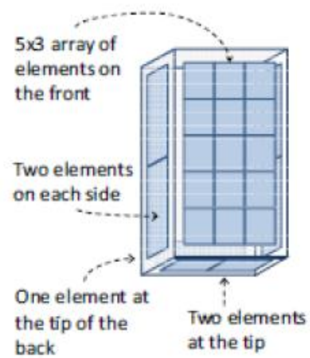
    - Is it closed? – *can move it aggressively*
    - Is it open and full? – *need to move with constraints*
    - Is it open and empty? – *move faster*

Hardware

PR2

Gripper – single dof parallel jaw

Fingertip sensors



Experiments

- Given a set of containers, determine if they are
  - Closed and full
  - Closed and empty
  - Open and full
  - Open and empty

Results

Given a feature vector and class, predict the internal state of the object.

- Dependency on probing force (Odwalla bottles)

f	*Recognition rate*
17 N	69.8 %
20 N	83.3 %
23 N	94.8 %

Human Study

- 17 human test subjects
  - Odwalla containers hidden from them, not allowed to move or pickup bottles, wearing gloves
  - Only allowed to use thumb and index finger
  - Noise-canceling headphones
  - Overall recognition rate – 72.2 %

- Feedback from most successful subjects
  - High-frequency feedback (tapping their fingers on the bottle)
  - Total compression and rate at which bottle returns to its original shape

Imitation Learning

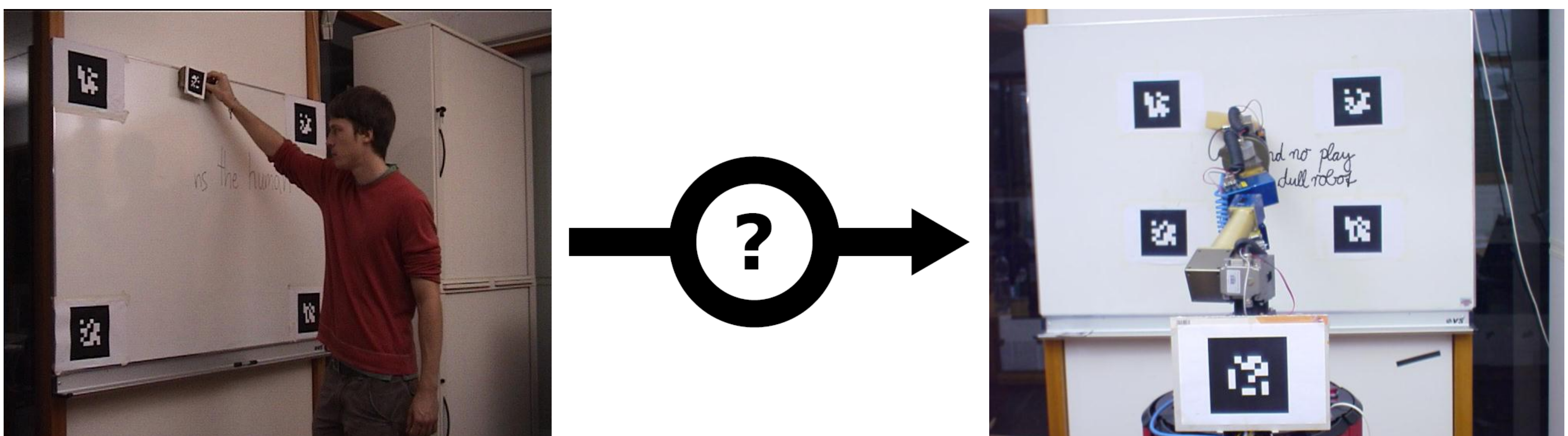
Problem Formulation

- **Given:**

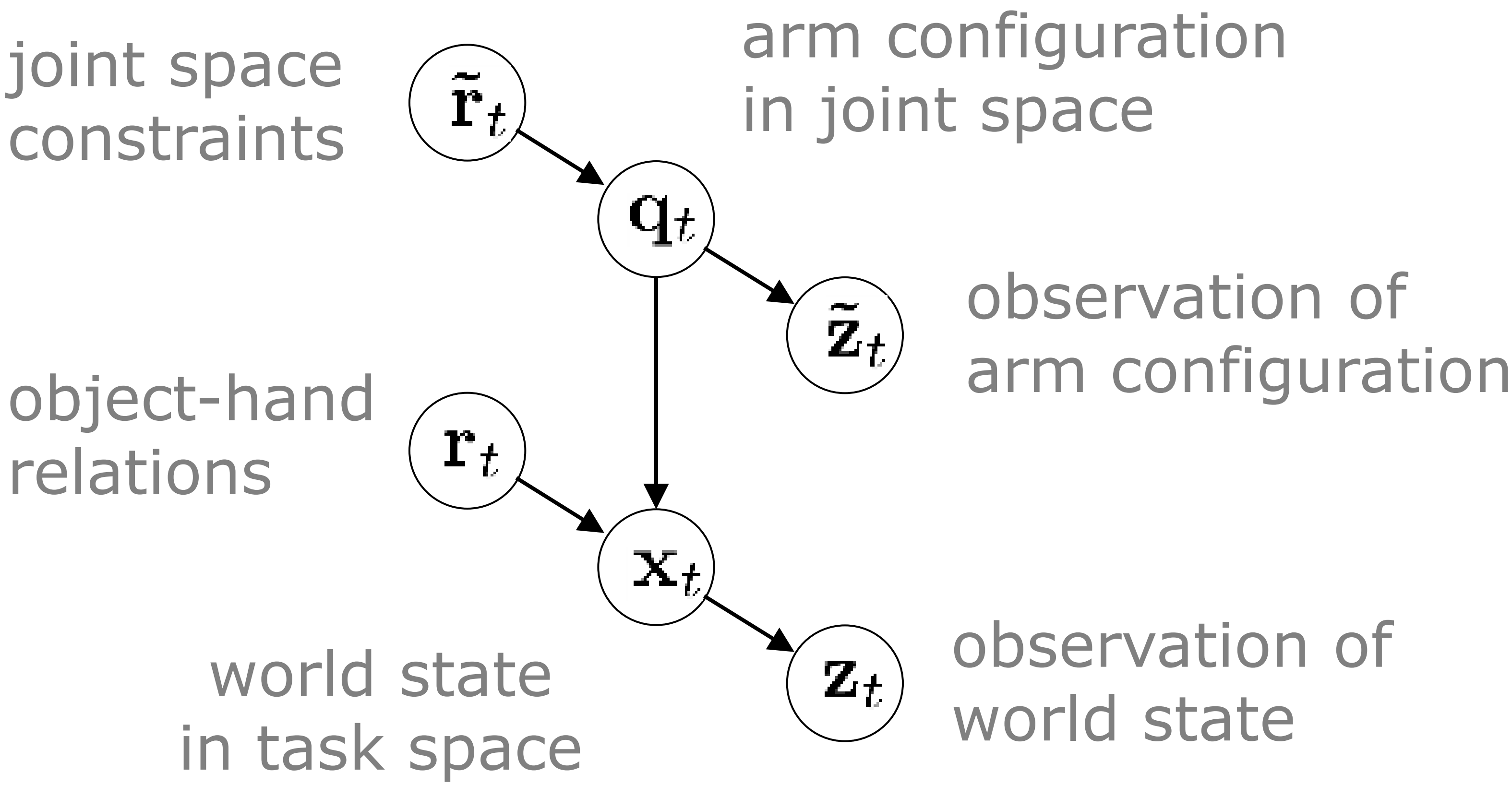
  Multiple demonstrations of the same manipulative task by a human teacher

- **Wanted:**

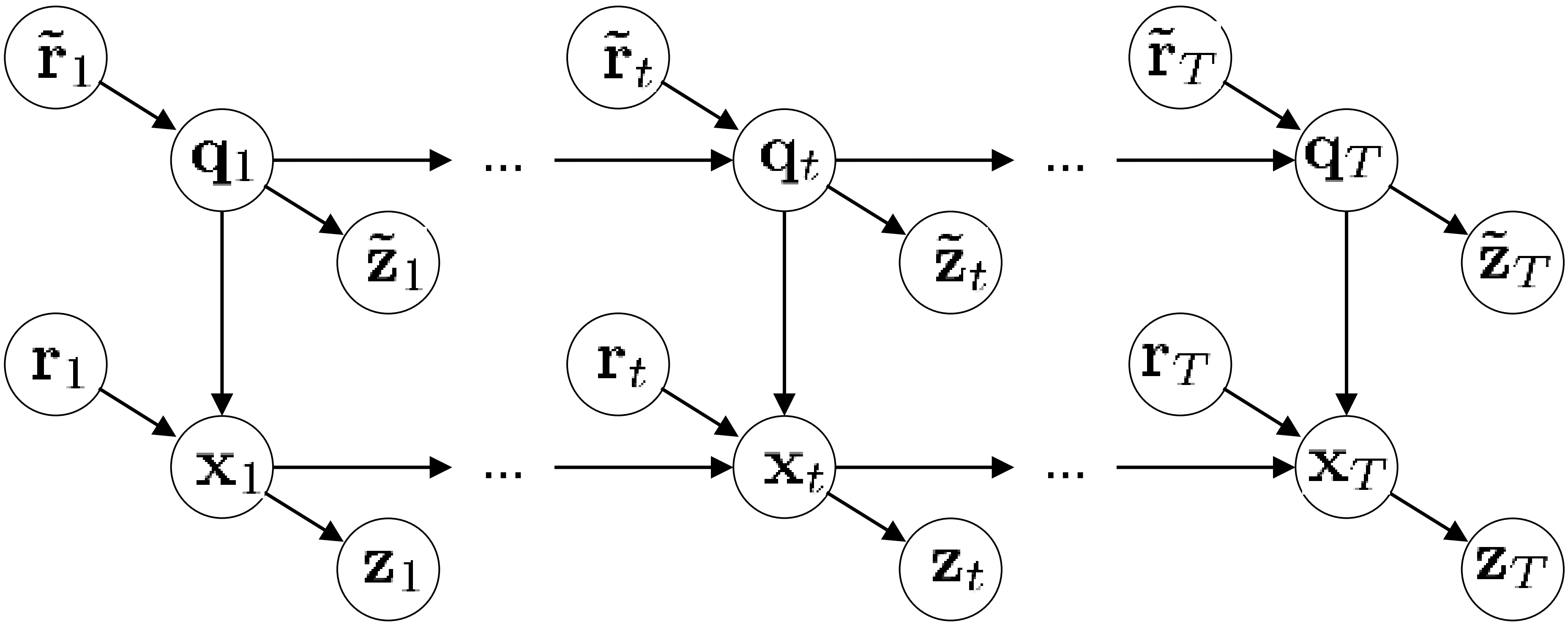
  A generalizable reproduction of the skill by a robotic manipulator



Dynamic Bayes Network for Imitation Learning

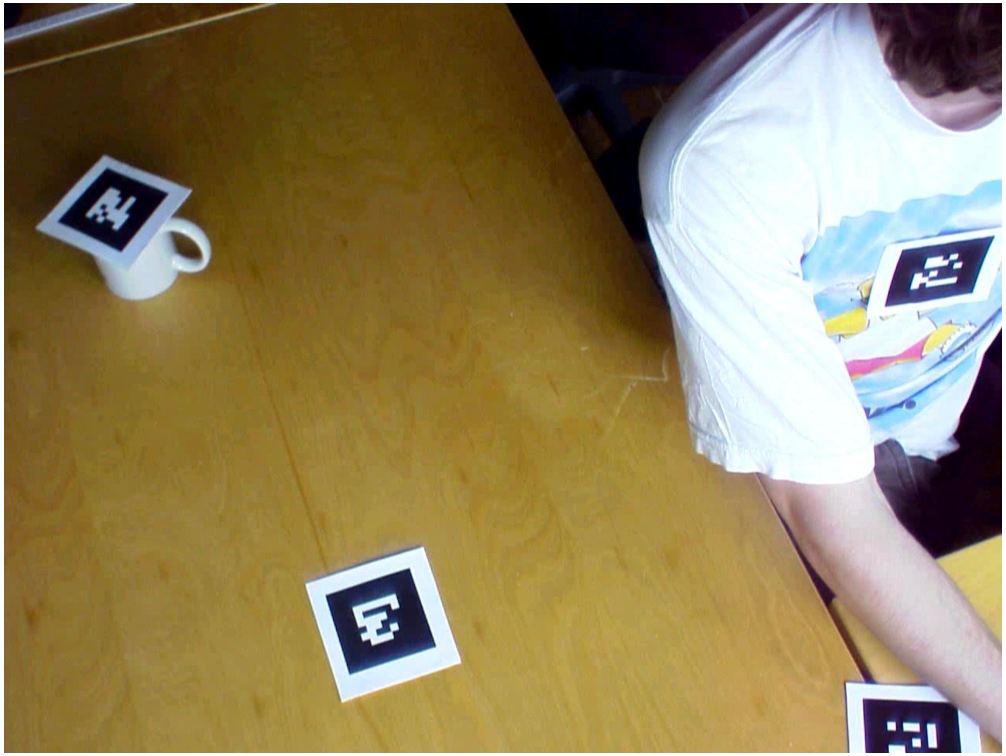


Dynamic Bayes Network for Imitation Learning



Task 1: Pick & Place (1)

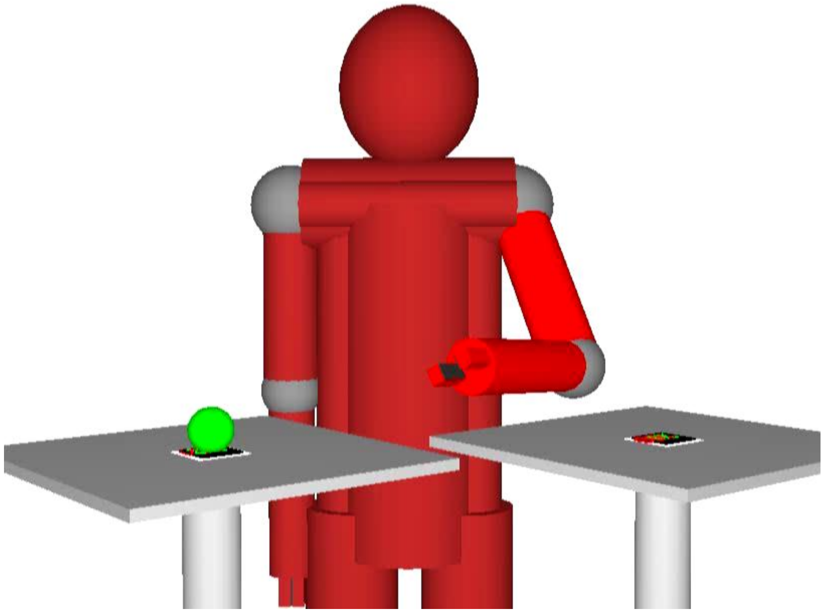
- Human demonstration



- Task: Pick cup and place on marker

Task 1: Pick & Place (2)

- Remove joint constraints

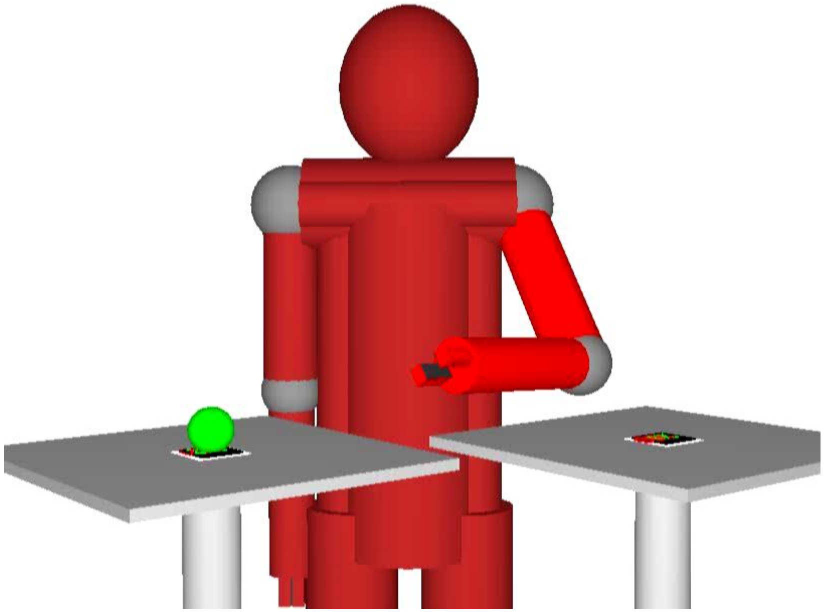


- Task learned successfully

➡ BUT: looks unnatural

Task 1: Pick & Place (3)

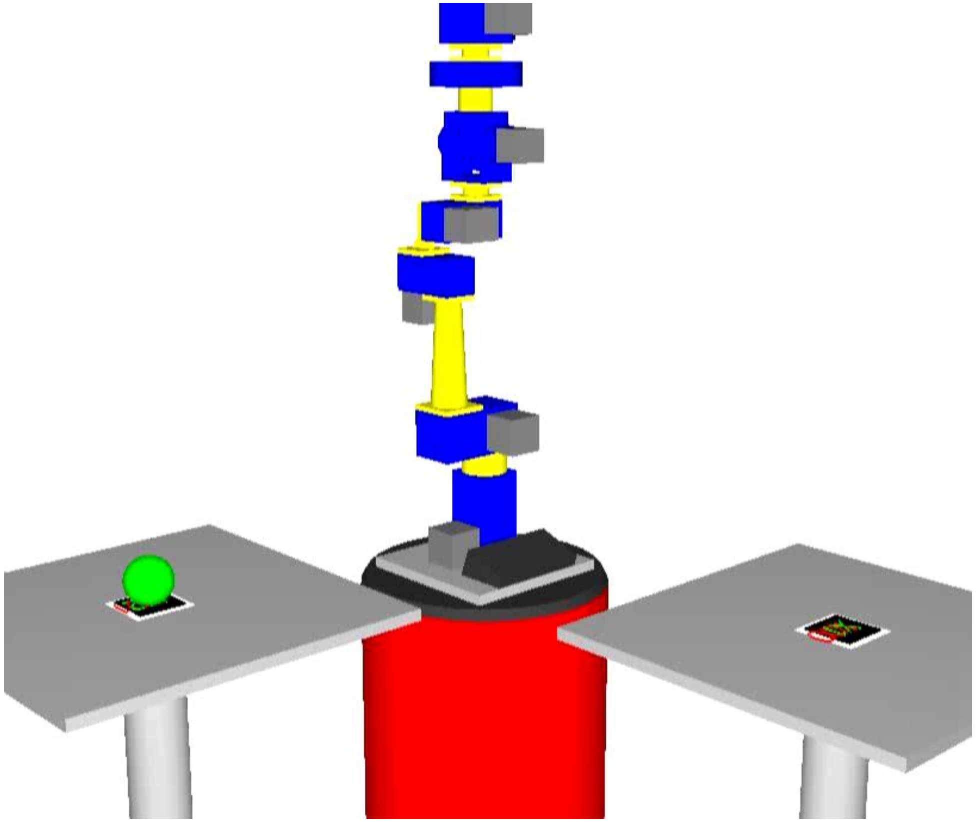
- With (learned) joint constraints



➡ Human-like movement

Task 1: Pick & Place (4)

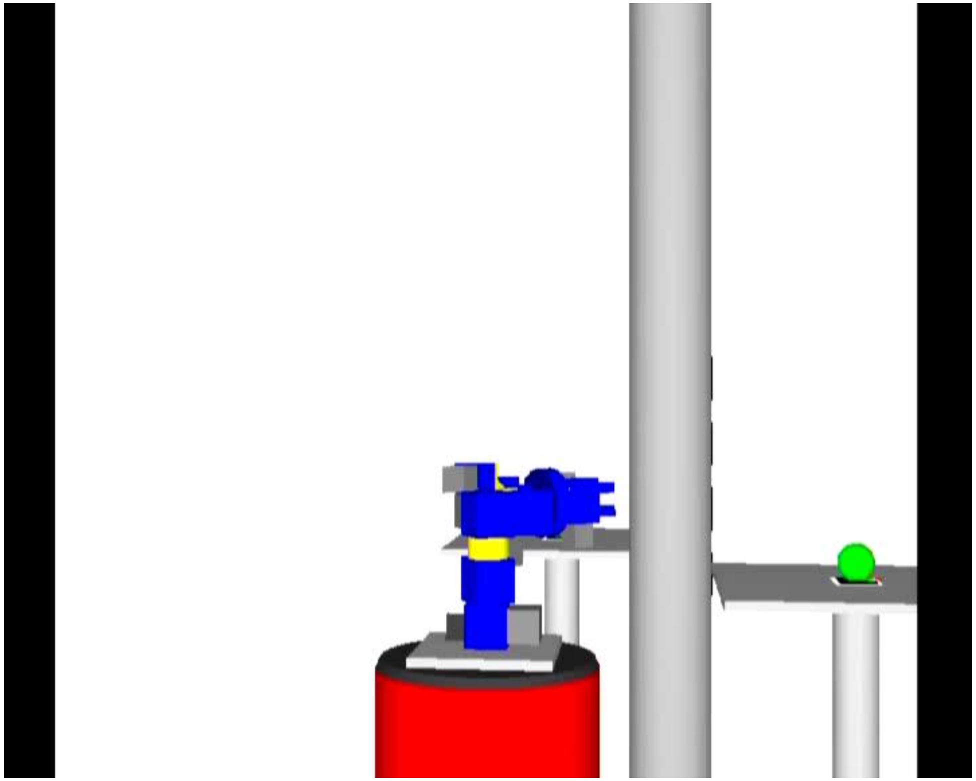
- Replace kinematic function



➡ Task is reproduced well

Task 1: Pick & Place (5)

- Add constraint for obstacle avoidance



➡ Task is reproduced well

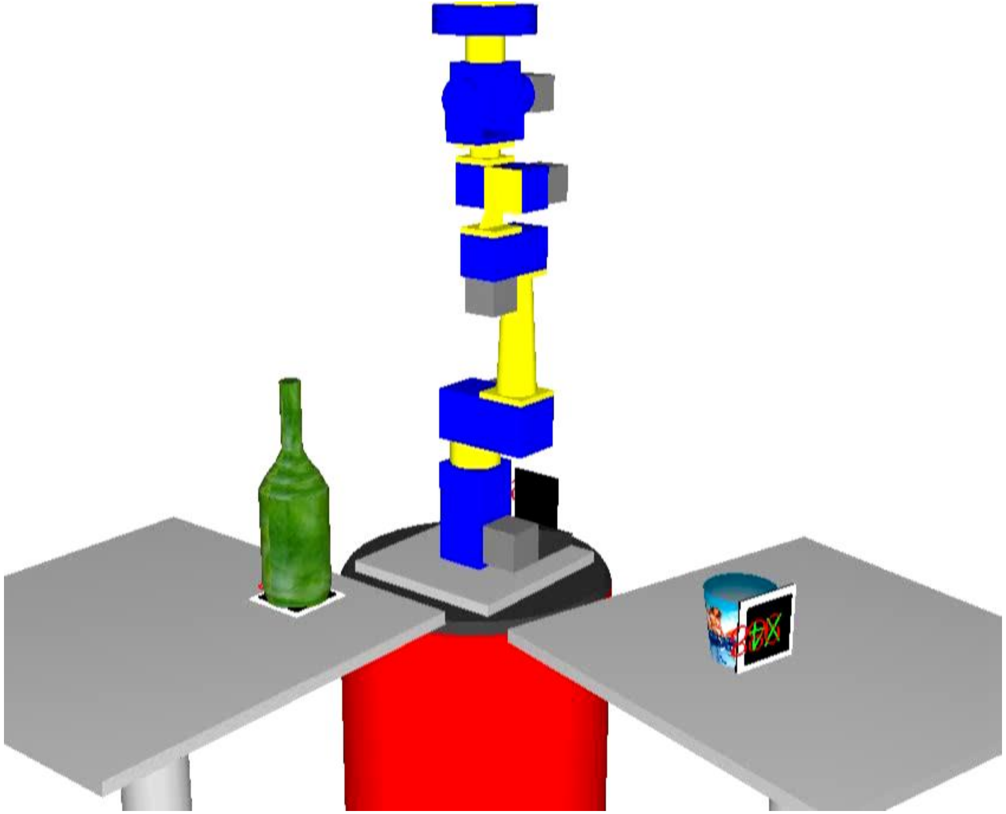
Task 2: Pouring (1)

- Human demonstration

➡ Extend state to include orientation

Task 2: Pouring (2)

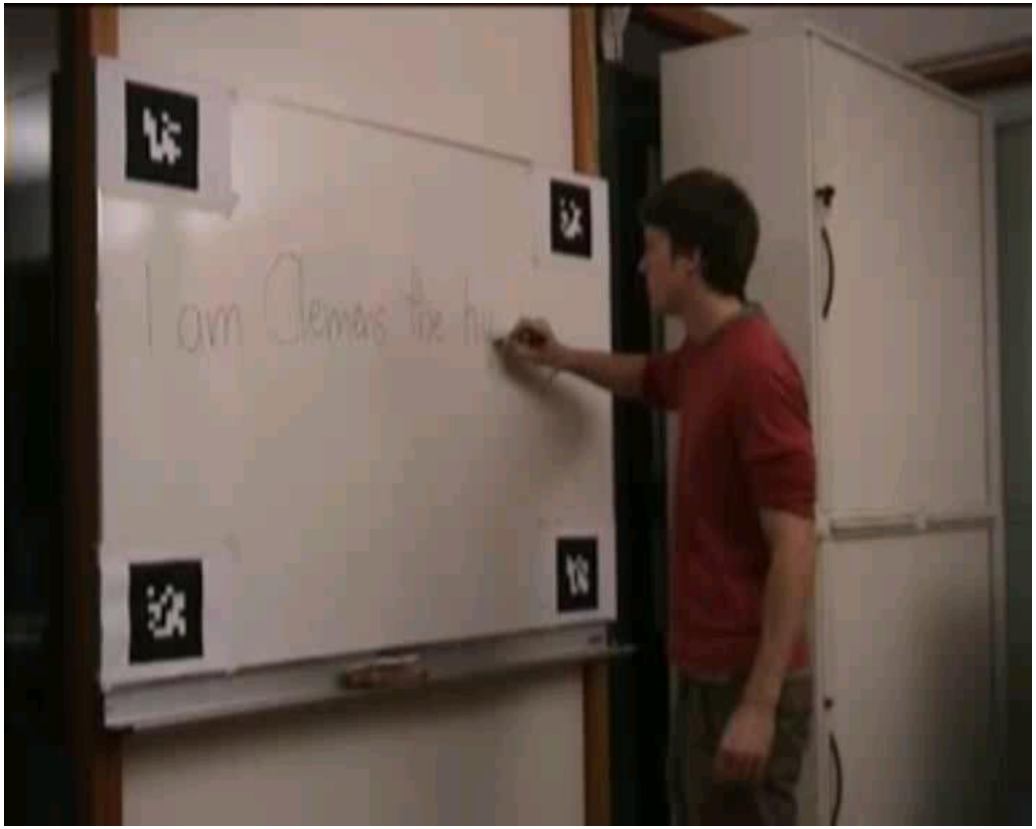
- Reproduction



➔ Task is well reproduced with 6D poses

Task 3: Whiteboard Cleaning

- Human demonstration



Task 3: Whiteboard Cleaning

- Robotic reproduction with obstacle